CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Analysis of Prediction Markets in Crypto: Investigating Convergence in Time, Volatility, and Biases in Polymarket

Master's thesis

Author: Bc. Pavel Řežábek
Study program: Economics and Finance
Supervisor: prof. PhDr. Ladislav Krištoufek, Ph.D.
Year of defense: 2024

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 25, 2024

Pavel Řežábek

Abstract

This thesis examines the dynamics of prediction markets within the cryptocurrency landscape, focusing on Polymarket. It offers a comprehensive analysis of these markets, exploring their functionality, potential for predicting future events, and volatility. The study highlights cognitive biases such as overestimating low probabilities and acquiescence bias, which influence market predictions. Volatility analysis reveals higher risks in prediction markets compared to traditional financial instruments, emphasizing the need for advanced risk management strategies. By addressing these biases and volatility, the research enhances understanding in behavioral finance, aiding traders in making informed decisions for more accurate market predictions.

JEL Classification	F12, F23, H25, H71, H87
Keywords	Prediction markets, Cryptocurrency, Polymarket, Financial instruments, Market volatility, Cognitive biases, Behavioral finance, Risk management, Market predictions, Overconfidence bias, Loss aversion, Gambler's fallacy, Sunk cost fallacy, Confirmation bias, Empirical analysis
Title	Analysis of Prediction Markets in Crypto: Investigating Convergence in Time, Volatility, and Biases in Polymarket

Abstrakt

Tato práce se zabývá dynamikou predikčních trhů v prostředí kryptoměn se zaměřením na Polymarket. Nabízí komplexní analýzu těchto trhů, zkoumá jejich funkčnost, potenciál pro předpovídání budoucích událostí a volatilitu. Studie poukazuje na kognitivní zkreslení, jako je přeceňování nízkých pravděpodobností a souhlasné zkreslení, které ovlivňují předpovědi na trhu. Analýza volatility odhaluje vyšší rizika na predikčních trzích ve srovnání s tradičními finančními nástroji, což zdůrazňuje potřebu pokročilých strategií řízení rizik. Tím, že se výzkum zabývá těmito zkresleními a volatilitou, zlepšuje porozumění v oblasti behaviorálních financí a pomáhá obchodníkům přijímat informovaná rozhodnutí pro přesnější předpovědi trhu.

Klasifikace JEL	F12, F23, H25, H71, H87
Klíčová slova	Předpovědní trhy, Kryptoměny, Polymarket, Finanční nástroje, Tržní volatilita, Kognitivní zkreslení, Behaviorální finance, Řízení rizik, Předpovědi trhu, Empirická analýza
Název práce	Analýza predikčních trhů v kryptoměnách: zkoumání konvergence v čase, volatility a zkreslení na Polymarketu

Acknowledgments

Author expresses profound gratitude to prof. PhDr. Ladislav Krištoufek, Ph.D.for his invaluable guidance, and extends special thanks to Prof. PhDr. Tomáš Havránek, Ph.D. and PhDr. Jiří Schwarz, Ph.D., for their insightful contributions, consultations and support.

Typeset in $\LaTeX$using the IES Thesis Template.

Bibliographic Record
Řežábek, Pavel: *Analysis of Prediction Markets in Crypto: Investigating Convergence in Time, Volatility, and Biases in Polymarket.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 70. Advisor: prof. PhDr. Ladislav Krištoufek, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

BTC Bitcoin

IEM Iowa Electronic Market

PoW Proof of Work

PoS Proof of Stake

EVM Ethereum Virtual Machine

dApps decentralized applications

GDP gross domestic product

USDC USD Coin

RMSE Root Mean Squared Error

UMA Universal Market Access - Oracle

CPMM constant product market maker

FPMM Fixed Product Market Maker

PM PolyMarket

WLS Weighted Least Squares

Yes token tradable asset representing a bet on an event's affirmative outcome

No token tradable asset representing a bet on an event's negative outcome

Chapter 1

Introduction

In the ever-evolving world of finance, the emergence of cryptocurrencies has introduced new economic possibilities and challenges. Among the various innovations within this space, decentralized prediction markets were introduced. These markets allow participants to trade contracts whose payout depends on the outcomes of future events, aggregating diverse information and opinions to forecast these events' outcomes. For example, one can bet on an event's outcome, such as the winner of a football match. As new information becomes available or the odds change, bettors can exit their bet position, realizing a profit or loss based on the current market rate, even before the event happens. This is how prediction markets work. This thesis aims to explore the operational mechanisms of decentralized prediction markets, examine their predictive potential using market efficiency measurements and time convergence, and measure and compare their volatility.

The primary objective of this thesis is to provide a comprehensive analysis of cryptocurrency-based prediction markets, with a particular focus on Polymarket. By examining these markets' operational mechanisms and unique characteristics, we aim to gain insights into their effectiveness in forecasting events and their behavior under different conditions. This analysis will help illuminate how these markets differ from traditional financial markets and the implications of these differences for traders and researchers alike.

Chapter 2 provides an in-depth overview of prediction markets, starting with a general definition and classification of these markets. It then explores the specifics of Polymarket, highlighting its operational principles and distinguishing features. Understanding these fundamentals is crucial for appreciating the complex dynamics that govern these markets.

Chapter 3 is dedicated to an extensive literature review. This chapter synthesizes existing research and scholarly articles related to prediction markets, both in the context of cryptocurrencies and more broadly. By drawing on a diverse range of sources, this chapter aims to construct a thorough background for the research in the next chapters and situates this study within the broader academic discourse.

Chapter 4 examines the convergence of probabilities over time within prediction markets. This chapter explores how predicted probabilities evolve as more information becomes available and the event gets closer in time. By analyzing the convergence patterns, we aim to understand the dynamics of the Polymarket better. As the outcome of an event approaches, predictions should become more accurate. In this section, we have empirically confirmed this hypothesis using data from Polymarket. Our analysis shows that predictions made five days before the decision are significantly more accurate than those made thirty days prior. Additionally, predictions made thirty days ahead are more accurate than those made ninety days in advance. This empirical evidence supports the notion that the accuracy of predictions in decentralized prediction markets improves as the event gets closer in time.

Chapter 5 focuses on the volatility observed in Polymarket. Volatility is a crucial measure of risk and uncertainty in financial markets. This chapter compares the volatility of Polymarket tokens with other financial instruments such as stocks and cryptocurrencies. The findings reveal that Polymarket is significantly more volatile than both stocks and cryptocurrencies. It also investigates how volatility varies across different timeframes and market types within Polymarket, providing valuable insights for traders to manage risk more effectively.

Chapter 6 explores the cognitive biases that affect prediction markets. Markets are not always efficient, and participants effective behavior. This chapter identifies and analyzes various cognitive biases, such as overconfidence, loss aversion, gambler's fallacy, sunk cost fallacy, confirmation bias, and acquiescence bias. The analysis demonstrates that both acquiescence bias and the overestimation of low probabilities are present in Polymarket. By understanding these biases, we can gain insights into the inefficiencies within prediction markets and develop strategies to mitigate their impact.

Finally, Chapter 7 summarizes the key findings of this study and discusses their implications for traders and researchers. It also outlines potential avenues for future research, emphasizing the need for continued exploration of prediction markets in the cryptocurrency domain and beyond.

In summary, this thesis endeavors to provide a comprehensive understanding of prediction markets in the cryptocurrency domain, with a focus on their operational principles, predictive potential, volatility, and the cognitive biases that influence them. By shedding light on these aspects, we aim to not only contribute to the broader field of decision sciences and behavioral finance but also to enhance the practical utility of prediction markets for various stakeholders. Our findings will equip traders and researchers with valuable insights, enabling them to navigate these markets more effectively and make informed decisions.

Chapter 2

Prediction markets and Polymarket

2.1 Prediction markets

A prediction market, also known as an information market, decision market, or event derivatives, as defined by Wolfers & Zitzewitz (2004a), is a platform where participants trade contracts whose payoff depends on the outcome of uncertain future events. These markets are designed to aggregate information and yield predictions about future events based on the collective wisdom and information of the participants.

Prediction markets are used in various fields, including finance, economics, politics, and public policy, to forecast events and trends more accurately than traditional methods. The theory is that market prices reflect the sum of all available information and thus provide more accurate forecasts than any individual expert could.

A prediction market differs from a betting broker by who is the counterparty to the trades - with a betting broker, it is the betting broker itself, whereas in a prediction market, the prediction market is just the intermediary, and the counterparty is another bettor.

Types of Prediction markets

One possible classification of prediction markets introduced by Zhan *et al.* (2021) is based on their operation mechanism. We distinguish between centralized and decentralized prediction markets.

Centralized Prediction Markets: A prominent example is the Iowa Electronic Market (IEM), which is widely recognized and utilized within the scientific community. This platform exemplifies the centralized prediction market model, where a single entity controls market operations. In the case of IEM, the central entity is the College of Business at the University of Iowa. Other instances include various corporate prediction markets, where companies create internal markets to forecast business-related outcomes.

Decentralized Prediction Markets: Examples of this type include Polymarket and Augur. These platforms operate on blockchain technology, ensuring not only decentralization but also a high level of transparency in their operations. However, in terms of complexity and features, these markets are generally considered more primitive when compared to well-established centralized markets like the Iowa Electronic Market. The blockchain-based markets are evolving, offering unique advantages like increased security and reduced reliance on central authorities, but they are still developing in terms of the sophistication and range of options available in more established markets.

Types of Prediction market contracts

Based on Oettler (2021) and others, prediction markets use several types of contracts to facilitate trading based on forecasts of future events. The most common types are:

- Binary Contracts: These are the simplest prediction market contracts. They have a yes/no format and payout a fixed amount if a specific event occurs, for example: "Will it rain in New York on July 4th?" If the event happens, the contract pays out; if not, it expires worthless. This is the most common contract on prediction markets and currently also the only possible on Polymarket

- Index-Based Contracts: These contracts payout based on a numerical value related to an event, such as the percentage of votes a candidate receives in an election or the closing value of a stock index on a particular day. For example: "You will get the same amount of dollars as how many percent of the population would vote for Republicans.".

- Spread Contracts: Similar to options in financial markets, spread contracts in prediction markets have a range of outcomes. Traders can buy and sell contracts based on whether the actual outcome will be above or below a certain point within that range.

- Futures Contracts: These are agreements to pay out based on the value of a variable at a future date. For instance, a contract might pay out based on the gross domestic product (GDP) growth rate of a country at the end of a quarter.

- Multi-Outcome Contracts: These contracts allow for more than two possible outcomes. For example, in a political election with multiple candidates, a multi-outcome contract could be created, and the bettor can then bet for any candidate, with payouts depending on who wins.

- Conditional Contracts: These contracts' outcomes depend on the occurrence of another event. For example, a contract might pay out only if a certain candidate wins an election and if they implement a specific policy within their first year in office.

Each type of contract is designed to capture different kinds of information and predictions, allowing participants to express their beliefs and insights in various ways.

Utilization of Prediction markets

Prediction markets have emerged as powerful tools for aggregating information and forecasting outcomes. These markets operate by allowing participants to buy and sell contracts based on their predictions about future events. The prices of these contracts reflect the collective wisdom of the participants, often leading to highly accurate predictions.

Research by Wolfers & Zitzewitz (2004b) has demonstrated the efficacy of prediction markets in political forecasting, often outperforming traditional polling methods. Arrow *et al.* (2008) highlighted their usefulness in economic and financial forecasting, noting that prediction markets can provide more accurate and timely predictions than individual experts.

In the corporate sector, Cowgill *et al.* (2009a) explored the application of prediction markets within organizations, finding that they can improve forecasts for project completion times and sales figures. Spann & Skiera (2003) showed that businesses could use prediction markets for new product development and market research, offering a cost-effective way to gauge potential success.

Overall, prediction markets offer a promising approach to enhancing the accuracy of forecasts across various fields by leveraging collective intelligence.

Corporate Prediction Markets

Built upon the premise of market efficiency, certain organizations have established internal prediction markets with the primary objective of obtaining highly accurate forecasts for critical company-related events, including sales figures and project deadlines. These internal prediction markets have been detailed by Cowgill & Zitzewitz (2013), with illustrations drawn from prominent companies like Google, Ford, and Koch Industries. The empirical findings from these cases indicate that the implementation of internal prediction markets resulted in a noteworthy 25 % reduction in Root Mean Squared Error (RMSE), signifying a marked improvement in forecast precision. Moreover, these internal markets demonstrated an enhanced level of efficiency as they evolved. In their initial stages, these prediction markets exhibited some inherent cognitive biases, such as an overly optimistic outlook on forthcoming events. Notably, corporate prediction markets found utilization across a diverse range of market-making entities.

Conditional prediction markets

Article "Prediction Markets as Decision Support Systems" (Berg & Rietz 2003) examines conditional prediction markets, using examples primarily drawn from the 1996 Presidential election. The authors leverage data from the Iowa Electronic Markets to explore the predictive capabilities of these markets, specifically examining the likelihood of each political party winning the elections under various candidate scenarios. The study defines conditional prediction markets as specialized platforms centered around forecasting future events conditional on specific conditions. The illustrative examples from the 1996 election highlight the novel concept of tying contracts to particular conditions. The research posits that both prediction and conditional prediction markets have broader applications beyond elections, serving as invaluable tools for decision support. The article suggests potential extensions into diverse domains such as movie marketing, business decision-making, and government policy evaluation, underscoring the versatility and relevance of prediction markets in various contexts.

2.2 Polymarket

One of the objectives of this thesis is to provide the reader with a comprehensive understanding of Polymarket operational mechanisms. It is crucial to understand these principles not only for the sake of comprehension but also to understand the distinctions that set Polymarket apart from other conventional markets. This foundational knowledge serves as a prerequisite for a more profound grasp of the empirical aspects that will be subsequently examined.

To facilitate a coherent understanding, we will begin by introducing and delineating the various fundamental concepts that form the backbone of the polymarkets functionality. These individual concepts, in turn, will be carefully placed within the broader context of market dynamics, creating a cohesive framework.

Blockchain

In the context of this master thesis, it is imperative to dive into the intricacies of blockchain technology, as it is the cornerstone of decentralized prediction markets. Blockchain, fundamentally a distributed ledger system, is distinguished by its decentralized framework and its capacity for immutable recordkeeping. This technology first emerged with the advent of Bitcoin (BTC), marking the first application of blockchain (Nakamoto 2008). At its core, blockchain operates by methodically sequentially recording financial transactions. Each transaction is encapsulated within a block, and these blocks are subsequently interconnected, forming a chain that represents a chronological and unalterable data record.

Central to the operation of blockchain is the implementation of a "consensus mechanism." This mechanism plays a pivotal role in determining how each new block is created and added to the chain, ensuring that the blocks are arranged in an orderly and sequential manner, and prevents conflict about who and how generates the next block. This aspect of blockchain is crucial for maintaining the integrity and reliability of the ledger. Within the realm of blockchain technology, the most notable consensus mechanisms are Proof of Work (PoW) and Proof of Stake (PoS). These mechanisms are not only essential for the addition of new blocks to the blockchain but are also fundamental in ensuring the security and decentralized nature of the technology.

EVM

A critical concept to understand is EVM, the Ethereum Virtual Machine, a critical component in most modern blockchain ecosystems, as introduced by Buterin (2013). The EVM sets a standard for the interpretation of orders written on the blockchain. It acts as a virtual environment where smart contracts and decentralized applications (dApps) are executed, providing a layer of abstraction between the executing code and the executing machine. This standardization is crucial for ensuring that applications behave consistently across different blockchain implementations.

The Ethereum Virtual Machine (EVM) can be conceptualized as a mechanism for interpreting and executing commands. Consider the scenario where a command is written into the blockchain, such as:

```
if (last_block_time > 1.1.2020):
    release 1 ETH from address_1 to address_2.
```

Once encoded and included in the blockchain, this command will be executed by the EVM. Consequently, the EVM facilitates the programming of transactions in a Turing-complete manner, enabling the creation of complex financial applications. The capability of the EVM to process complex conditional statements and logic makes it a powerful tool for developing decentralized applications that can manage sophisticated financial operations.

A practical illustration of the EVM's application is Polymarket, a prediction market that operates on the Polygon blockchain. Polygon is noteworthy for its compatibility with the EVM, which allows it to leverage the robust framework and widespread adoption of Ethereum-based applications while offering improvements such as enhanced scalability and lower transaction costs. Within Polymarket, individual markets and trades are essentially programmed orders on the blockchain. These orders are executed in the context of the EVM, allowing for complex financial interactions to be made in a decentralized, secure, and transparent manner.

The integration of platforms such as Polymarket on EVM-compatible blockchains like Polygon highlights the versatility and efficiency of the EVM standard. It enables a seamless interpretation of blockchain orders, increasing an environment where innovative financial solutions can thrive. This aspect of blockchain technology, particularly the role of the EVM, is a testament to the evolving landscape of decentralized applications and their potential impact on various sectors. Building Polymarket on top of an EVM-compatible blockchain ensures its decentralization, as this is an inherent blockchain property. This approach guarantees robustness and stability, leveraging the underlying blockchain's architecture to provide a secure and resilient platform. By utilizing an EVM-compatible blockchain, Popymarket inherits the advantages of decentralized ledger technology, enhancing the system's overall integrity and reliability. Additionally, the compatibility with other blockchains and the ease of auditing the code, given that EVM is an industry standard, further strengthen Popymarket's framework.

Tokenization

Tokenization is a pivotal concept within blockchain technology crucial also for prediction markets. Token is a digital representation of a specific asset or utility.

It involves the conversion of real-world assets into digital tokens, which are then securely recorded on a decentralized ledger. This transformative process relies on the fundamental principles of blockchain, including transparency, immutability, and decentralization (Underwood 2016). By representing real-world assets as digital tokens on a blockchain, tokenization introduces a new paradigm in asset management, offering plenty of benefits and applications across various industries. Tokenization also involves the generation of tokens with no relation to the real world. An example might be the creation of a new cryptocurrency token on top of other cryptocurrency blockchains, such as the Shiba-inu token.

Diverse types of tokens have emerged, each with its distinct characteristics and functions. Utility tokens, for instance, provide access to a specific product or service within a blockchain ecosystem, while security tokens represent ownership of an asset and may offer dividends. On the other hand, stablecoins are designed to minimize price volatility by pegging their value to a stable asset, often a fiat currency.

Tokenization has diverse advantages. Increased liquidity is a notable benefit, as digital tokens can be easily traded on various blockchain platforms. Accessibility is enhanced through fractional ownership, allowing individuals to invest in high-value assets with minimal capital. Transparency is increased, as every transaction is recorded on an immutable ledger, fostering trust among participants. Additionally, security improvements are achieved through cryptographic techniques that safeguard the integrity of the tokenized assets.

The underlying technology facilitating tokenization is blockchain, with its key components being smart contracts and decentralized ledgers (Tapscott & Tapscott 2017). Smart contracts execute predefined rules autonomously, enabling the automatic transfer of digital tokens based on predefined conditions. Decentralized ledgers ensure the integrity and transparency of the tokenized assets by distributing the record across a network of nodes.

Examining specific use cases, Polymarket stands out as a unique tokenization application. In this platform, every bet is represented by a yes-token and a no-token. Upon resolution, owners of the correct token (depending on the outcome) can exchange it for a USDC token. This demonstrates how tokenization can be applied in the context of prediction markets, introducing a novel way to engage in speculative activities.

In conclusion, tokenization represents a transformative element in the financial and asset management landscape. With its diverse types, inherent benefits, and applications across sectors, tokenization showcases the evolving potential of blockchain technology. Real-world examples, such as stablecoins, further underscore the practical implications and innovative possibilities that emerged from this paradigm shift in how we perceive and manage assets.

Stablecoin

Based on Berentsen & Schar (2019), stablecoin is a digital unit of value with the following three properties: 1) It is not a form of currency, 2) it can be used without any direct interaction with the issuer, 3) it is tradable on a secondary market and has a low price volatility in terms of a target quote currency.

In practice, a stablecoin operates as a convertible token that consistently maintains a value of one dollar.

In decentralized prediction markets like Polymarket, stablecoins, particularly USDC, play a crucial role in maintaining the value of bets, as the bets are denominated in the units of USDC. By using a stablecoin, users can focus solely on speculating the outcome of events without the concern of currency value fluctuations. USDC, the stablecoin used by Polymarket, ensures stability through off-chain collateral, keeping its value consistent and independent of typical cryptocurrency volatility. This design allows for a more predictable and straightforward betting environment.

The stability of the USDC coin is maintained through the issuer's commitment to redeem or issue these tokens at a fixed value of one dollar. In the event of a deviation from the $1 peg, arbitrage opportunities arise, attracting traders who engage in transactions that drive the price back to the targeted one-dollar value (Consortium 2021; Lipton & Hardjono 2020).

Oracle

In the intricate landscape of blockchain technology, integrating real-world information onto the immutable blockchain is facilitated through entities known as oracles. These oracles are pivotal in various blockchain applications, including decentralized prediction markets such as Polymarket. The necessity for oracles within Polymarket arises from their fundamental role in resolving market outcomes based on real-world data, a critical function in the operation of prediction markets (Xu *et al.* 2021).

An oracle functions as the source of information input to the blockchain. For example, if a transaction depends on weather conditions, such as whether it is raining, the oracle provides the blockchain with this information. The main challenge is to determine the truth in a decentralized manner, establish incentives for the oracle to provide accurate information, and implement quality correction mechanisms to address any false information (Zhang *et al.* 2020).

The oracle problem, often identified as a primary vulnerability in smart contracts, has become a focal point of discussion. This difficulty stems from ensuring the accuracy and reliability of real-world data that smart contracts rely on for execution. Different types of oracles have emerged, each presenting unique advantages and disadvantages. A classification that holds particular relevance in the context of Polymarket (Werner *et al.* 2021).

The three main types of oracles, classified according to their mechanisms, are as follows:

1. **Trusted Oracle**: A trusted oracle operates based on trust in a specific entity. If this entity provides false information, the outcome is based on incorrect data, and users are left without correction. The primary incentive for the oracle to provide truthful information is that an oracle known for dishonesty will lose credibility and usage over time, ultimately becoming obsolete.

2. **Optimistic Oracle**: An optimistic oracle proposes a result that is considered valid unless any party disputes it. If a dispute arises, the oracle transitions to a dispute resolution mechanism, often involving multiple oracles in the process.

3. **Multiple Oracle**: A multiple oracle system requires consensus among multiple entities to determine the truth from the outset. This system involves assigning different weights to each participating entity and establishing a clear process for reaching agreement.

By their very nature, decentralized prediction markets depend on oracles to inform about and verify real-world events, as these events dictate the resolution of market outcomes. The accuracy and trustworthiness of oracles become crucial to the integrity of the entire prediction market ecosystem, thereby addressing the oracle problem (Harz & Bano 2021).

In the specific case of Polymarket, a sophisticated approach to the oracle problem is employed through the utilization of UMA (Universal Market Access) as an optimistic oracle. This choice reflects a strategic decision to leverage a decentralized and optimistic model for obtaining real-world information, aligning with the principles of transparency and decentralization inherent in blockchain technology. It would not be so with the trusted oracle.

In essence, the interplay between oracles, particularly in Polymarket's use of UMA as an optimistic oracle, highlights the intricate dynamics involved in bringing real-world information onto the blockchain. As blockchain technology evolves, addressing the Oracle problem becomes crucial in increasing the robustness of smart contracts and decentralized applications.

Liquidity provision mechanism

Liquidity is a fundamental and indispensable element within any market, which is pivotal in ensuring its smooth and efficient operation. Liquidity allows for the fast execution of trades by providing a ready market for buying or selling assets. In a liquid market, assets can be traded without significant delays, ensuring that transactions are carried out efficiently. Adequate liquidity minimizes the impact of large trades on asset prices. In illiquid markets, substantial trades can lead to price slippage, causing the executed price to deviate significantly from the intended price. Liquidity acts as a stabilizing force by preventing drastic and unpredictable price fluctuations. In a liquid market, there is a continuous flow of buying and selling assets, which helps maintain stability and mitigate the risk of sharp market movements.

Polymarket uses liquidity pools for trading tokens, a standard method in blockchain trading. This process involves traders contributing assets to a liquidity pool, facilitating token exchange based on supply and demand. The liquidity pool contains Asset A and Asset B, and it is possible to exchange one asset for the other such that $A \times B = K$. K is a constant number. For the possibility to make an exchange with the pool, you need to contribute a fee to the pool. This liquidity pool type is also referred to as CPMM - constant product market maker (Adams *et al.* 2020). Alternatively, it is possible to put assets into the pool, and the depositor gets a share of future fees. But at the same time, he will be threatened by the so-called divergence loss caused by the exchange rate change (Eisenberg & Lehar 2020).

This mechanism works by creating a pool of tokens A and B. For instance, one can deposit 10 tokens of A and 10 tokens of B into the pool, resulting in a constant product K of 100. Participants can trade within this pool such that the product $A \times B$ remains equal to 100. For example, if a participant deposits 5 tokens of A, they can withdraw 3.33 tokens of B, ensuring that the product $A \times B$ continues to equal 100. This mechanism allows pool creators to earn profits on their tokens and grow their wealth (Angeris & Chitra 2020).

The next operation involves increasing liquidity in the pool. When additional assets are added to the pool, the contributor's share of the pool is determined by the increase in the square root of K. For instance, if the pool initially contains 10 tokens of A and 10 tokens of B, and an additional liquidity provider contributes 5 tokens of A and 5 tokens of B, the value of K increases from 100 to 225. Consequently, the square root of K rises from 10 to 15, giving the new liquidity provider a one-third share of the pool. This operation does not result in any gains or losses for the participants but does enhance the pool's liquidity. Conversely, participants can withdraw their liquidity under the same conditions (Adams *et al.* 2019).

However, the issue of pool liquidity becomes significant when the relative prices of tokens A and B change. This change alters the asset composition within the pool, leading to a decrease in overall value. In extreme cases, where the value of one token drops to zero, the pool becomes flooded with an infinite amount of the valueless token while the quantity of the valuable token approaches zero. This phenomenon is known as divergence loss (Feng *et al.* 2021).

Polymarket recently introduced a system similar to traditional stock exchanges. In this model, users can place put and call options, allowing any other participant to engage in these trades. This innovation offers several advantages, such as increased transparency in pricing, more direct control over trades, and potentially reduced slippage, leading to a more efficient and user-friendly trading experience (Polymarket 2023a).

Polymarket step-by-step

To mint tokens, a participant deposits 1 USDC into the market's "account", resulting in the creation of one "yes" token and one "no" token. Upon market resolution, this dollar is redeemed for either the “yes” token or the “no” token, depending on the outcome (Polymarket 2023a).

A Fixed Product Market Maker (FPMM) liquidity pool is established to facilitate the trading of “yes” and “no” tokens. Participants can add liquidity to the pool, earning a share of the pool and and trading fees in the process. The trading mechanism allows participants to exchange their “yes” tokens for “no” tokens and vice versa. The key principle governing these trades is that the product of the number of “yes” tokens and “no” tokens must equal a constant, K, which remains unchanged before and after a trade (Polymarket 2023b).

Individual participants can then engage in trading by acquiring new tokens. From a new user’s perspective, they are essentially “betting” on either “yes” or “no.” In the background, this action results in the minting of both “yes” and “no” tokens, which are subsequently exchanged in the pool for the tokens that correspond to the user’s bet (Polymarket 2023a).

Participants have the option to burn their tokens. By returning both “yes” and “no” tokens, they can reclaim the one dollar initially locked during the token minting process. This mechanism is useful for participants who prefer not to wait for the market resolution. Alternatively, tokens can be sold to other users if there is sufficient demand (Polymarket 2023a).

Technically, trades are never terminated but become irrelevant once the market outcome is apparent. Before that, participants withdraw their tokens from the liquidity pool to prevent the divergence loss. The oracle then records the result, and a dispute period follows to address any disagreements over the resolution. Historically, disputes have occurred in only 3 out of 187 markets. Finally, the winning tokens are redeemed for 1 USDC each, corresponding to the original amount locked during token minting (Polymarket 2023a).

Chapter 3

Literature review

This literature review explores the theoretical foundations, empirical evidence, and factors influencing the convergence of prediction market prices to true probabilities, the efficiency of prediction markets, and the volatility observed in these markets compared to traditional financial assets.

Hanson (2003) introduced the concept of market scoring rules, which underpin many prediction markets. These rules facilitate the updating of prices as new information becomes available, theoretically guiding prices towards the true probabilities. The notion of information aggregation in prediction markets is further elaborated by Manski (2006a), who discusses how market prices can be interpreted as probabilities under certain conditions.

Empirical studies have investigated the accuracy and convergence properties of prediction markets across various domains. Berg *et al.* (2003) conducted a comprehensive analysis of the Iowa Electronic Markets (IEM) and found that market prices tend to converge to the true election outcomes as the election date approaches. Similarly, Forsythe *et al.* (1992) examined presidential elections and observed that prediction markets provide more accurate forecasts than traditional polls.

Research by Snowberg *et al.* (2012) highlighted that prediction market prices adjust rapidly to new information, demonstrating the markets' efficiency in information processing. Moreover, Tetlock & Mellers (2015) found that prediction markets often outperform expert judgment, particularly in scenarios with high uncertainty.

Several factors influence the rate and accuracy of convergence in prediction markets. Market liquidity, defined by the volume of trades and number of participants, plays a crucial role in ensuring that prices reflect true probabilities

(Ritholtz 2008). Spann & Skiera (2009) found that markets with higher liquidity tend to exhibit faster and more accurate convergence. The diversity of information among participants is another critical factor. Surowiecki (2005) argues that diverse and independent information sources contribute to better aggregate predictions. Ottaviani & Sorensen (2009) support this view, noting that prediction markets with heterogeneous participants often achieve more accurate outcomes.

The convergence of prediction market prices to true probabilities is supported by both theoretical and empirical evidence. While market liquidity and participant diversity are essential for accurate predictions, the overall effectiveness of prediction markets in forecasting remains robust across various domains. Future research could further explore the impact of different market designs and external factors on prediction accuracy.

The efficiency of prediction markets is a fundamental concept that correlates the betting rate with the probability of the event. Fama (1970) introduced three types of market efficiency, each delineating the degree to which information is reflected in asset prices. Weak-form efficiency posits that prices instantly and fully reflect all past price information, making future price movements unpredictable based on past prices. Semi-strong efficiency suggests that asset prices fully reflect all publicly available information, allowing only investors with inside information to gain an advantage. Strong-form efficiency asserts that asset prices fully reflect all public and inside information, ensuring no one can have an advantage in predicting prices.

Confirming efficiency is inherently elusive and subject to potential refutation due to statistical significance and future unpredictability. The statistical principle of significance indicates that every 20th random variable is expected to be significant at a 95 % confidence level, complicating the definitive establishment of efficiency. Furthermore, the unpredictability of future events means that even if a variable predicted prices in the past, it might not do so in the future, as another participant may have identified and capitalized on this inefficiency, restoring market efficiency.

Manski (2004) challenges traditional approaches in predicting choice behavior through revealed preference analysis and rational expectations, advocating for measuring expectations through subjective probabilities using survey research. Manski's work highlights the limitations of traditional methods and explores the history and emerging literature on eliciting probabilistic expectations, suggesting this approach could validate assumptions about expectations and enhance empirical foundations for decision-making analysis across various domains, including macroeconomic events, risks, future income, and individual choices.

Addressing the skepticism raised by Manski (2004), Wolfers & Zitzewitz (2006) establishes analytic foundations for interpreting prediction market prices as corresponding with mean beliefs, outlining sufficient conditions for this correspondence. They demonstrate that prediction market prices usually align with the mean beliefs of traders, providing useful estimates of average beliefs about event probabilities and offering a micro foundation for the claim that prediction markets efficiently aggregate beliefs. Manski (2006b) discusses the formal analysis of price determination in prediction markets with diverse beliefs, revealing that the equilibrium price is a quantile of the budget-weighted distribution of beliefs, does not disclose the mean belief but establishes a bound, and remains unchanged even if traders adjust their beliefs using price data. This analysis underscores the caution needed when interpreting prices as market probabilities, particularly in real prediction markets.

Prediction market efficiency can be compromised by biases influencing trader behavior and market outcomes. Acquiescence bias, a tendency for respondents to agree with statements regardless of their actual beliefs, can significantly distort prediction market results if participants conform to prevailing sentiments rather than expressing true expectations (Knowles & Nathan 2005). Atkinson (2012) examined how the phrasing of questions and market structure can lead to overestimation of certain events' probabilities, suggesting careful consideration of wording and format to mitigate acquiescence bias. Similarly, the overestimation of low probabilities bias, also known as probability weighting, where rare events are overestimated and common events underestimated, impacts market efficiency. Kahneman & Tversky (1979a) identified this bias, with Gonzalez & Wu (1999) and Barberis (2013) proposing methods to correct it, including adjusting market prices based on historical data and implementing sophisticated trading algorithms.

Prediction market volatility is another critical aspect for forecasting events. Prediction markets, such as Polymarket, a leading decentralized platform, have gained attention for forecasting various events. These markets allow participants to trade contracts based on future events' outcomes, efficiently aggregating information and providing accurate predictions. Research has shown prediction markets' effectiveness in domains like politics and economics (Wolfers & Zitzewitz 2004b), with volatility influencing their reliability and stability.

Comparing prediction market volatility with traditional financial markets like stocks and cryptocurrencies offers insights into their stability and risk. Stocks and cryptocurrencies exhibit varying degrees of volatility influenced by market sentiment, regulatory changes, and macroeconomic factors (Cont 2001; Urquhart 2016). Team (2021) found that prediction markets have different volatility patterns, driven by event-specific information rather than broader economic trends, suggesting unique advantages and risks compared to traditional assets.

The temporal aspect of prediction market volatility is crucial for understanding their behavior leading up to events. Tetlock *et al.* (2008) indicates that prediction market prices stabilize as the event approaches and more information becomes available, contrasting with financial markets where volatility can increase due to speculative trading and external shocks. Different event types exhibit varying levels of volatility. Political events, for example, often see heightened volatility during elections due to rapidly changing information and voter sentiment (Rothschild 2009). In contrast, markets predicting economic indicators might display more stable volatility patterns influenced by scheduled data releases and established economic models (Lewis 2011).

Understanding prediction market volatility, particularly in comparison to traditional financial assets and across different event types, is essential for assessing their effectiveness and reliability as forecasting tools. Polymarket, as a leading platform in this domain, offers a unique perspective on how decentralized prediction markets operate. Future research should continue to explore these dynamics, incorporating more granular data and advanced modeling techniques to enhance our understanding of prediction market volatility.

Chapter 4

Convergence of probability in time

4.1 Introduction

Understanding the behavior of predicted probabilities over time is essential in the context of prediction markets.

In probability and prediction, the dynamic nature of how probabilities evolve over time is a fascinating subject. As events unfold, the probabilities associated with their outcomes do not remain static; instead, they tend to shift and adjust as more information becomes available. This chapter focuses on the concept of probability convergence over time, particularly within the framework of prediction markets.

Prediction markets are unique in their structure, as they revolve around the binary nature of event outcomes: either an event occurs, or it does not. Unlike traditional financial markets where asset values fluctuate indefinitely, prediction markets converge to a definitive outcome: either 1, indicating the event has occurred, or 0, indicating it has not. This convergence process is influenced by several critical factors that will be explored in depth in this chapter.

Figure 4.1 shows an example of predicted probabilities in time in a given prediction market.

By understanding how and why these predicted probabilities align more closely with actual outcomes as the event horizon narrows, we can gain insights into market behavior and improve predictive accuracy. This chapter will provide a comprehensive analysis of the convergence patterns, comparing different market types and their respective speeds of convergence, ultimately aiding traders in making more informed and strategic decisions.

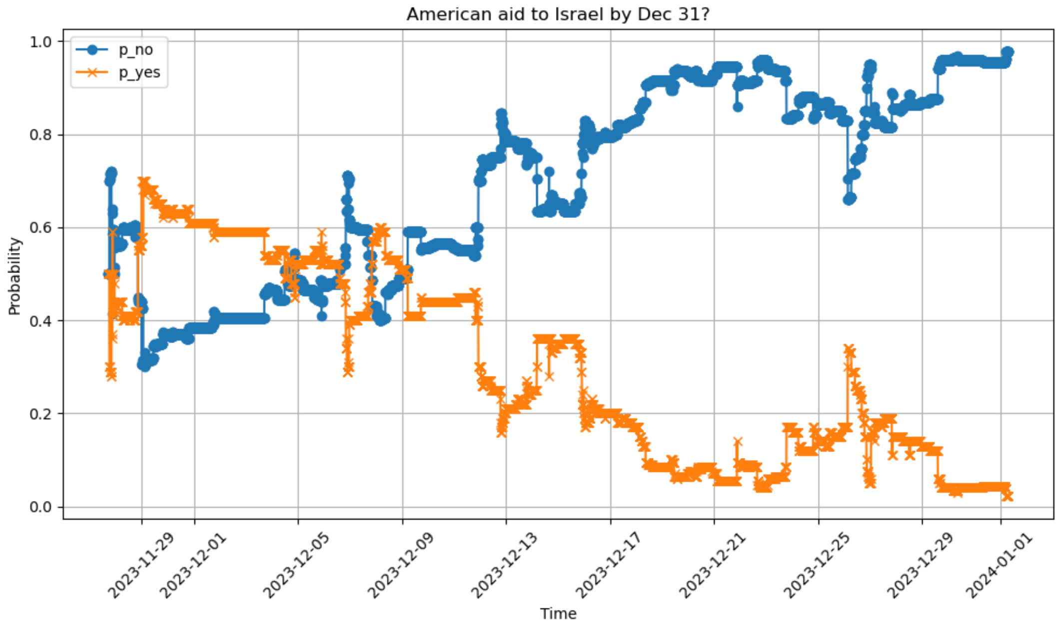

Figure 4.1: Example Prediction Market: Evolution of Predicted Probability Over Time

Once an event occurs, its probability becomes a certainty (i.e., the probability of the outcome is 1). Conversely, before the event happens, the probability can range between 0 and 1. Our research demonstrates that the probability of the final outcome tends to increase over time due to three primary factors:

Firstly, Increasing Information Availability: As time progresses, information accumulation enhances predictions' accuracy. Continuous generation and assimilation of information into the market refine the predicted probabilities.

Secondly, Diminishing Time for New Information: As the event nears, the opportunity for new, impactful information to emerge decreases. This reduction in the arrival of new information further stabilizes the predicted probabilities.

Thirdly, Obvious Outcomes: In certain scenarios, the outcome becomes apparent well before the event concludes, leading to early stabilization of the predicted probability.

This pattern of gradual convergence of predicted probability to the final outcome over time is consistent across various market types. In this sense, prediction markets are very different from other capital markets, as, for example, stocks do not converge to a specific value. In prediction markets, the final token value is either 1 or 0, reflecting the occurrence of the given event.

This chapter empirically confirms the tendency of predicted probabilities to converge toward the final outcome over time. It also compares the convergence trends between different market types, addressing whether specific market types tend to converge faster than others. Figure 4.2 illustrates the temporal development of average predicted probabilities, providing a visual representation of how predictions converge as the resolution date approaches.

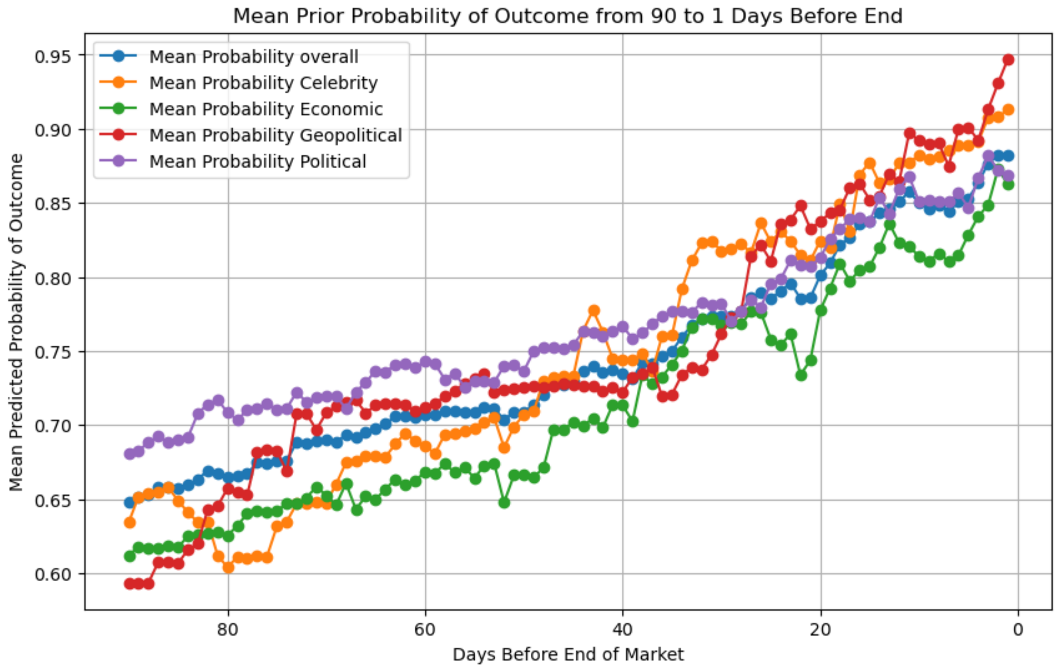

Figure 4.2: Convergence of Probability Over Time: Graph Showing Average Predicted Probability Evolution 90-0 days in advance, Separated by Market Types

Understanding these convergence trends is crucial for traders, as it allows them to better estimate expected returns when betting on the true outcome. By gaining insights into the speed and pattern of convergence, traders can make more informed decisions, enhancing their ability to capitalize on prediction markets.

4.2 Data and Methodology

This section will comprehensively detail the data sources and methodologies employed in our research. By precisely outlining our approach, we aim to provide a clear and thorough explanation of the methods upon which our study is built.

Market convergence

Two methods were employed to empirically confirm that prediction markets systematically converge to the final outcome. Firstly, a linear model was prepared. The model was specified as:

$$\text{The predicted probability of final outcome} \sim \text{days before resolution}$$

Data used for this regression were: predicted probability of an outcome for each day from 90 days in advance up to 0 days in advance and number of days in advance. Data contains all markets resolved between 31.12.2023 and 1.4.2024. One of the main linear model assumptions is that there is no autocorrelation (independence of observations). As this time series data are autocorrelated from their essence, the model leads to biased interpretations, and its results serve only to point out the relationship.

In order to make a robust statistical measurement, we took the predicted probability 90, 60, and 5 days before the final bet evaluation for each market. We subtracted these probabilities from each other. Then, using a student t-test, we measured whether the difference in predictions of the correct outcome over time was statistically significant.

The assumptions of one sample t-test are Independence and Normality of Variance. Independence is satisfied in this case because the individual markets are independent of each other at one-time point.

The normality assumption seems to be satisfied because the distribution looks approximately normally distributed.

The assumption of homogeneity of variance does not hold in our scenario, as we use a one-sample t-test to compare the difference between predicted probabilities across different time frames, specifically comparing them to zero.

Categories comparison

Our research aimed to determine if specific categories of prediction markets tend to converge to the final outcome earlier in the next measurement.

Prediction markets were systematically categorized into four distinct groups based on their thematic content:

1. Celebrity Actions: This category includes markets predicting actions by celebrities, such as the prediction market bettingn on of Taylor Swift receiving a marriage proposal.

2. Economic Indicators: This category encompasses markets related to eco-

nomic metrics, such as the probability of Bitcoin reaching its all-time high by January 31.

3. Geopolitical Events: This category covers markets focused on geopolitical developments, such as the market betting on if of a peace deal between Saudi Arabia and the Yemeni Houthis by the end of 2023.

4. Political Events: This category includes markets predicting political occurrences, such as the probability of Mark Cuban announcing a presidential run by the end of the year.

We compared each category 90, 30, and 5 days in advance. Our analysis satisfied the assumptions necessary for conducting the two-sample t-tests. Firstly, the data appeared to be normally distributed upon visual inspection. Secondly, tests for heteroskedasticity did not reveal any significant evidence of unequal variances, thus meeting the assumption of homogeneity of variances. Lastly, the observations in our dataset are inherently independent of each other, ensuring that the assumption of independence is upheld.

4.3 Results and discusion

Market convergence The results of our t-test analysis indicate statistically significant differences in the predicted probabilities over time. Specifically, the average difference between the probabilities 90 and 30 days before the final bet evaluation is 12,6 percent points with a standard deviation of 21,3 percent points. Similarly, the average difference between 30 and 5 days is 7,9 percent points with a standard deviation of 20 percent points as shown in table 4.2.

	Coef	Std.Err.	t
Intercept	0.860	0.003	264
day	0.003	< 0.001	40

Table 4.1: Convergence in Time: Linear Model of Daily Predicted Probability Approach to the Final Result

The results of our t-test analysis indicate statistically significant differences in the predicted probabilities over time. Specifically, the average difference between the probabilities 90 and 30 days before the final bet evaluation is 12,6 percent points with a standard deviation of 21,3 percent points. Similarly, the average difference between 30 and 5 days is 7,9 percent points with a standard deviation of 20 percent points as shown in table 4.2.

	90-30	**30-5**
Average	0.126	0.079
Standard Deviation	0.213	0.200

Table 4.2: Differences in Predicted Probabilities of Outcome Between Different Advance Times: Descriptive Statistics

The t-test results, as presented in table 4.3 further support the significance of these differences, with a t-statistic of 8.0 (p-value $= 6.3 \times 10^{-14}$) for the 90 to 30 day comparison, and a t-statistic of 5.36 (p-value $= 1.2 \times 10^{-7}$) for the 30 to 5 day comparison. Both p-values are well below the conventional threshold of 0.05, indicating that the differences in prediction accuracy are statistically significant.

	90-30	**30-5**
t-statistic	8.01	5.36
p-value (one-sided)	$\ll 0.0001$	$\ll 0.0001$

Table 4.3: Differences in Predicted Probabilities of Outcome Between Different Advance Times: T-Test Results

The strong statistical significance of the t-test results (p-values far below 0.05) provides robust evidence that the markets are indeed becoming more accurate in their predictions as the event date nears.

This finding supports the hypothesis that prediction markets are aggregating information and improving accuracy over time, ultimately converging towards the true outcome. Thus, our study statistically validates the theory that markets converge in probability in time to the true outcome.

Categories comparison

Tables 4.4 and 4.6 show the average probability of outcomes in each market category 90 and 5 days respectively. Tables 4.5 and 4.7 show comparison statistics.

When the different market types are compared 90 days in advance, only one has a p-value less than 0.05, indicating significantly different rates of approaching the actual impact. This suggests that political issues are easier to predict in prediction markets 90 days in advance.

When the different market types are compared five days in advance, no p-values are less than 0.05, indicating no significantly different rates of approaching the actual impact among the market categories in the short term. This implies that prediction markets converge similarly across different categories as the event approaches.

In conclusion, the analysis of prediction markets over different time horizons reveals interesting insights. While political outcomes show significantly better predictability 90 days in advance compared to other categories, this distinction diminishes as the event date approaches. Five days in advance, the prediction accuracy across all market categories becomes statistically indistinguishable, highlighting the balance in the system. This convergence suggests that as more information becomes available closer to the event, the predictability of different market categories equalizes, underscoring the dynamic nature of prediction markets and their ability to assimilate information effectively over time.

Category	Avg. Probability of Outcome	St. Deviation
Celebrity Actions	63.4 %	20.9 %
Economic Indicators	61.2 %	20.6 %
Geopolitical Developments	59.3 %	18.9 %
Political Outcomes	68.1 %	21.2 %

Table 4.4: Average Probability of Outcome 90 Days in Advance: Descriptive Statistics of Various Prediction Market types

Comparison	t-statistic	p-value
Celebrity Actions vs. Economic Indicators	0.50	0.61
Celebrity Actions vs. Geopolitical Developments	0.96	0.34
Celebrity Actions vs. Political Outcomes	-1.05	0.30
Economic Indicators vs. Geopolitical Developments	0.44	0.66
Economic Indicators vs. Political Outcomes	-1.56	0.12
Geopolitical Developments vs. Political Outcomes	-2.06	0.04

Table 4.5: Probability of Outcome 90 Days in Advance: T-Test Comparison Between Various Prediction Market Types

Category	Avg. Probability of Outcome	St. Deviation
Celebrity Actions	88.9 %	15.4 %
Economic Indicators	82.9 %	19.3 %
Geopolitical Developments	90.1 %	14.9 %
Political Outcomes	84.7 %	21.3 %

Table 4.6: Average Probability of Outcome 5 Days in Advance: Descriptive Statistics of Various Prediction Market types

Comparison	t-statistic	p-value
Celebrity Actions vs. Economic Indicators	1.63	0.11
Celebrity Actions vs. Geopolitical Developments	-0.39	0.70
Celebrity Actions vs. Political Outcomes	1.06	0.29
Economic Indicators vs. Geopolitical Developments	-1.99	0.05
Economic Indicators vs. Political Outcomes	-0.43	0.67
Geopolitical Developments vs. Political Outcomes	1.39	0.17

Table 4.7: Average Probability of Outcome 5 Days in Advance: Comparison Statistics 5 Days in Advance

Chapter 5

Polymarket volatility

5.1 Introduction

Volatility is a fundamental measure of risk and uncertainty in financial markets. In this chapter, we undertake a comprehensive analysis of the volatility observed in Polymarket, a decentralized prediction market platform. The motivation for this chapter is to equip traders with valuable insights into volatility, enhancing their risk management strategies. Traders can make more informed decisions by understanding how volatility varies across different timeframes and market types. A deep dive into volatility is particularly important for Polymarket, as it exhibits significantly higher volatility than traditional markets such as the stock market. Several key factors might contribute to the higher volatility in Polymarket:

- Limited Liquidity Relative to Average Bet Size: Insufficient market liquidity means that any substantial bet can significantly impact prices. The larger the bet, the more pronounced the price movement.
- High Sensitivity to New Information: The market demonstrates acute sensitivity to new information, resulting in substantial volatility as fresh data is continuously integrated.
- Market Manipulation: The intentional manipulation of prices or dissemination of false information by certain participants is a notable concern. Due to the relatively low liquidity and lack of stringent regulation in the crypto market, a single large bettor can deliberately induce significant price swings.

This chapter will examine volatility from three perspectives:

1. Comparative Analysis with Traditional Markets: To provide context, we will compare Polymarket volatility with the volatility of stocks and bonds.

2. Temporal Volatility Analysis: We will analyze volatility across different time periods within Polymarket. According to the study by Heimbach *et al.* (2023), liquidity providers often withdraw from liquidity pools before the end of the market to avoid divergence loss. Additionally, the frequency of new information tends to increase as the resolution date approaches. Therefore, we hypothesize that volatility is higher closer to the resolution date than at earlier stages.

3. Market Type Comparison: We will comprehensively compare volatility between different types of markets within Polymarket. This analysis aims to identify patterns and differences in price fluctuations across various market categories, which can provide valuable insights for traders. Understanding these volatility dynamics is crucial as it can help traders make informed decisions about where to allocate their investments, manage risks more effectively, and optimize their trading strategies based on the specific characteristics of each market type.

In the subsequent sections, we will outline the data utilized for this analysis and the methodology employed. The final section will present the results and their interpretation.

5.2 Data and Methodology

In examining prediction markets, which inherently operate as zero-sum entities where the average return is zero, we employed a methodology centered on volatility analysis. Volatility was quantified using the standard deviation of daily returns for the respective “yes” and “no” tokens within each market.

For comparative purposes, we analyzed volatility across several asset classes, including major cryptocurrencies and stocks. The specific steps involved are as follows:

1. **Data Collection**:

    - **Polymarket Data**: Price data for all Polymarket markets ending between 31.12.2023 and 31.3.2024 were collected.

- **Cryptocurrency Data**: Price data for major cryptocurrencies (BTC, ETH, XRP, ADA, MAT, BCH, XMR) were gathered for approximately the same period as Polymarket data, from October 1, 2023, to March 31, 2024. Note that the Polymarket data starts around September 1, as it is approximately 90 days before ending on 31.12.2023.
- **Stock Market Data**: Volatility data for the SP 500 stocks were also collected for the same period from October 1, 2023, to March 31, 2024.

2. **Volatility Calculation:**

   The daily volatility of each asset class (Polymarket markets, cryptocurrencies, and stocks) was calculated using the standard deviation of daily returns.

3. **Comparative Analysis Using Welch T-test:**

   We used the Welch T-test to compare the volatility of prediction markets with that of cryptocurrencies and stocks. The Welch T-test was chosen over the standard Student's T-test due to its violation of the equal variances assumption. The data appeared normal, fulfilling the T-test's normality assumption. The independence of observations was evident in this case.

4. **Analysis of Polymarket volatility in time:**

   We computed the average volatility and the standard deviation of volatility across three distinct time periods preceding the event: 90-61 days in advance, 60-31 days in advance, and 30-1 days in advance. Subsequently, we employed a T-test to assess whether any of these time periods exhibited significantly higher volatility compared to others. This statistical approach enabled us to determine if the observed variations in volatility were significant and consistent across the defined time frames. The assumptions of the standard Student's T-test (homogeneity of variance, normality, and independence of observations) appeared to be fulfilled in this case, as our initial analyses indicated that the volatility data met the necessary conditions. Specifically, the data exhibited similar variances across the time periods, followed a normal distribution, and the observations were independent as they were drawn from non-overlapping intervals.

5. **Volatility Comparison Across Different Types of Prediction Markets**:

   We calculated the volatility of each Polymarket market and used the standard Student's T-test to compare the volatility of different types of prediction markets. The assumptions of homogeneity of variance, normality, and independence of observations appeared to be fulfilled in this case, justifying the use of the standard Student's T-test.

The subsequent sections will detail the results of this analysis and provide an interpretation of the findings. Through this examination, we aim to offer traders enhanced insights into market volatility, thereby aiding in the development of more effective risk management strategies.

5.3 Results and interpretation

Comparison of volatility with other market instruments

Comparison	t-value	Degrees of Freedom
Cryptocurrencies vs Stocks	6.1	5.4
Prediction Markets vs Stocks	25.1	188.5
Prediction Markets vs Cryptocurrencies	4.6	6.9

Table 5.1: Volatility Comparison of Polymarket and Other Assets: Welch's T-Test Results

Table 5.1 provides Welch's T-test results of a statistical comparison of the volatilities between stocks, major cryptocurrencies, and prediction markets. These tests were conducted to determine if the differences in volatility between these asset classes are statistically significant at the 0.05 significance level ($\alpha = 0.05$). Below is the interpretation of each comparison.

Stocks vs. Prediction Markets

The critical value for a two-tailed test with $\alpha = 0.05$ and approximately 189 degrees of freedom is around ± 1.972. The absolute t-value 25.1 far exceeds 1.972, demonstrating that the volatility of stocks is significantly lower than that of prediction markets. This finding highlights the extreme volatility present in prediction markets, indicating a much higher level of risk compared to traditional stock markets.

Cryptocurrencies vs. Prediction Markets

The critical value for a two-tailed test with $\alpha = 0.05$ and approximately 7 degrees of freedom is around ± 2.365. The absolute t-value 4.6 is significantly greater than 2.365, suggesting that the difference in volatility between cryptocurrencies and prediction markets is statistically significant. This result illustrates that, despite the high volatility in cryptocurrencies, prediction markets exhibit even greater volatility, emphasizing the substantial risk involved.

Conclusion

The Welch's T-test analysis confirms that prediction markets exhibit significantly higher volatility compared to both stocks and cryptocurrencies at the 0.05 significance level. This higher volatility suggests increased risk and instability in prediction markets, which traders must consider when engaging in these markets. The statistical significance of these results provides robust evidence of the varying levels of risk associated with different asset classes, guiding traders in their risk management strategies.

Comparison Analysis of Volatility in Time

To determine if there are statistically significant differences in volatility across three defined time periods, we conducted a t-test. The primary objective was to assess whether any specific time period exhibited higher volatility compared to others.

Timeframe	Mean	Standard Deviation
1 (90-61 days in advance)	0.073	0.057
2 (60-31 days in advance)	0.058	0.055
3 (30-1 days in advance)	0.059	0.063

Table 5.2: Polymarket Volatility Comparison Over Time: Descriptive Statistics of Volatility Across Different Timeframes

Comparison	t-value	p-value
Period 1 (90-61 days) vs. Period 2 (60-31 days)	1.039	0.151
Period 1 (90-61 days) vs. Period 3 (30-1 days)	0.925	0.179
Period 2 (60-31 days) vs. Period 3 (30-1 days)	-0.045	0.518

Table 5.3: Polymarket Volatility Comparison Over Time: Descriptive Statistics of Volatility Across Different Timeframes

The results of the t-tests as presented in table 5.3 are as follows:

Comparison between Time Period 1 and Time Period 2:

The t-value of 1.039 and p-value of 0.151 suggest that there is no statistically significant difference in volatility between Time Period 1 and Time Period 2. Since the p-value is greater than the common significance level of 0.05, we fail to reject the null hypothesis.

Comparison between Time Period 1 and Time Period 3:

The t-value of 0.925 and p-value of 0.179 indicate that there is no statistically significant difference in volatility between Time Period 1 and Time Period 3. The p-value exceeds 0.05, leading us to fail to reject the null hypothesis. We, therefore, can not confirm that the volatility tends to increase as the event approaches in time. On the contrary, volatility seems to be declining over time.

Comparison between Time Period 2 and Time Period 3:

The t-value of -0.045 and p-value of 0.518 suggest that there is no statistically significant difference in volatility between Time Period 2 and Time Period 3. The p-value is much greater than 0.05, reinforcing the decision to fail to reject the null hypothesis.

Conclusion

The one-sided t-tests provide valuable insights into the temporal dynamics of market volatility. Specifically, our analysis indicates that there are no statistically significant differences in volatility across the three defined time periods (as shown in 5.2). This suggests that volatility does not significantly change as the event date approaches. These findings imply that, within the observed data, market participants do not exhibit increased uncertainty or instability closer to the event date. Further research could explore the underlying causes of these volatility patterns and their implications for market behavior and prediction accuracy.

Volatility Comparison Across Different Types of Prediction Markets

The analysis of volatility across different types of prediction markets as shown it tables 5.5 and 5.4 reveals the following insights based on the Student's T-test results:

1. **Economic Indicators vs. All Except Economic Indicators**: The t-value of 2.637 and p-value of 0.009 indicate a statistically significant difference in volatility between economic indicators and all other categories combined. The higher volatility in economic indicators suggests that markets predicting economic outcomes are more volatile, likely due to the complexity and variability of economic factors.

Category	Mean Volatility	St. Deviation
Economic Indicators	0.13	0.06
Political Outcomes	0.11	0.07
Celebrity Actions	0.10	0.04
Geopolitical Developments	0.10	0.06
Excl. Economic Indicators	0.10	0.06
Excl. Political Outcomes	0.11	0.06
Excl. Celebrity Actions	0.11	0.06
Excl. Geopolitical Developments	0.11	0.06

Table 5.4: Polymarket Volatility Comparison Across Different Market Types: Descriptive Statistics

Comparison	t-value	p-value
Economic Indicators vs. Others	2.63	0.01
Political Outcomes vs. Others	-1.49	0.14
Celebrity Actions vs. Others	-2.16	0.03
Geopolitical Developments vs. Others	-1.84	0.07

Table 5.5: Polymarket volatility comparison across different market types: T-Test Results

2. **Political Outcomes vs. All Except Political Outcomes**: The t-value of -1.49 and p-value of 0.14 suggest no statistically significant difference in volatility between political outcomes and the other categories combined. This indicates that the volatility observed in markets predicting political outcomes is comparable to that in other prediction markets.

3. **Celebrity Actions vs. All Except Celebrity Actions**: The t-value of -2.16 and p-value of 0.03 indicate a statistically significant difference in volatility between celebrity actions and all other categories combined. The lower volatility in celebrity actions suggests these markets are relatively stable compared to others, possibly due to the less complex and more predictable nature of celebrity-related events.

4. **Geopolitical Developments vs. All Except Geopolitical Developments**: The t-value of -1.84 and p-value of 0.07 suggest that there is no statistically significant difference in volatility between geopolitical developments and the other categories combined, although the p-value is close to the threshold for significance. This indicates a moderate level of volatility in geopolitical developments, reflecting the unpredictable nature of such events.

Conclusion

The Student's T-test analysis indicates significant differences in volatility across various types of prediction markets. Specifically, economic indicators and celebrity actions exhibit distinct volatility patterns compared to other categories, highlighting the unique risk profiles of these markets. Understanding these volatility patterns can help traders make more informed decisions based on the type of prediction market they are engaging with.

Chapter 6

Biases and prediction markets

6.1 Introduction

Understanding market efficiency is complex, as there is no straightforward way to measure it. Market inefficiency, by definition, indicates a deviation of prices from their true values. However, determining these true values is inherently challenging. If it were easy to find the correct value of an asset or bet, traders would not engage in transactions at prices differing from this value, resulting in an efficient market by default.

Betting markets present a unique case for studying market efficiency due to their inherent characteristics. Unlike stock markets, where the correct value of a stock is never conclusively known, betting markets eventually reveal the "correct" value of a bet, which resolves to either \$1 or \$0.

Biases in markets are often reflections of the biases held by individual participants. When these participants deviate from perfectly rational behavior, they exhibit cognitive biases. Such biases are ingrained in human psychology and influence decision-making in everyday life as well as in financial contexts. This chapter focuses on exploring and analyzing cognitive biases within prediction markets, specifically examining how these biases can impact market outcomes.

The primary objective of this chapter is to explore the presence and impact of cognitive biases in prediction markets. We will review the relevant biases, their nature, their function, and their implications for market outcomes. This exploration will be followed by a data-driven analysis to empirically measure some of the biases within prediction markets. By conducting this analysis, we aim to confirm the existence of biases and thereby challenge the notion of market efficiency in prediction markets.

Given that we have investigated Polymarket, a relatively new and small prediction market, it is reasonable to hypothesize that these markets may exhibit lower efficiency levels than more established markets. This chapter will identify and review various cognitive biases likely to influence prediction market participants. It will describe the data and analytical techniques used to identify and measure biases in prediction markets. Finally, the chapter will present and interpret the findings from the data analysis, highlighting the presence and impact of these biases.

Understanding these biases is crucial for optimizing trading strategies and gaining a deeper comprehension of market behavior. By shedding light on the cognitive biases affecting prediction markets, this chapter contributes to the broader field of behavioral finance. It enhances our understanding of market efficiency in this unique environment of decentralized prediction markets.

6.2 Identify Cognitive Biases

In this section, we discuss common cognitive biases and their potential effects on Polymarket.

The Overconfidence Effect

The overconfidence effect is a cognitive bias where an individual's subjective confidence in their judgments is reliably greater than their objective accuracy, especially when confidence is relatively high (Pompian 2006).

In prediction markets, the overconfidence effect significantly influences the decision-making of betting agents. This cognitive bias results in an inflated perception of one's ability to predict market trends, consequently leading to excessive risk-taking and the misallocation of bets. While it is highly probable that overconfidence impacts a substantial number of bets, the anonymity of bettors poses a challenge for quantifying this effect within the scope of this thesis.

Loss Aversion

Loss aversion is a principle in behavioral economics and cognitive psychology that asserts that losses typically have a more substantial psychological impact on individuals than equivalent gains. Simply put, people tend to prefer avoiding losses over acquiring gains (Kahneman & Tversky 1979b).

In prediction markets, the principle of loss aversion plays a crucial role in influencing decision-making behavior. Individuals experience the pleasure of winning less intensely than the pain of losing, meaning that the psychological distress caused by a loss is greater than the joy experienced from a gain of the same magnitude. Consequently, even if the expected value of a bet or investment is positive, the potential for loss can deter individuals from engaging in otherwise favorable opportunities. This bias can counteract the overconfidence effect, which drives individuals to take greater risks than are justified. While overconfidence can lead to excessive betting or investment, loss aversion can result in underinvestment or overly cautious behavior. Although the precise impact of loss aversion is challenging to quantify, it is a critical factor to consider when analyzing the behavior of agents in prediction markets. Understanding this bias helps in comprehending why individuals might act conservatively despite favorable odds or why they might fail to capitalize on positive expected values due to the disproportionate fear of losses.

Gambler's Fallacy

The Gambler's Fallacy is the erroneous belief that if a particular event occurs more frequently than normal during a given period, it will be less likely to happen in the future (or vice versa), despite the events being independent (Tversky & Kahneman 1971).

In prediction markets, the Gambler's Fallacy can lead to misguided decision-making by betting agents. Believing that recent outcomes will influence future events, agents may adjust their bets based on perceived patterns rather than actual probabilities. This misjudgment can result in poor strategic choices, distorting market efficiency and leading to suboptimal allocation of resources and capital. Such behavior undermines rational market operations, highlighting the critical impact of cognitive biases in economic environments.

Sunk Cost Fallacy

The sunk cost fallacy is a cognitive bias that leads individuals to continue an endeavor once an investment in money, effort, or time has been made, even when continuing is no longer the best course of action (Arkes & Blumer 1985).

In prediction markets, the sunk cost fallacy can drive betting agents to persist with losing bets or strategies due to prior investments, regardless of current market conditions or future prospects. This bias leads to inefficient decision-making, as agents may prioritize staying in opened loss positions over optimizing future gains. Consequently, resources are misallocated, and market dynamics are skewed, potentially reducing overall market efficiency and profitability. Understanding and mitigating the sunk cost fallacy is crucial for maintaining rational investment strategies and achieving optimal market outcomes.

Confirmation Bias

Confirmation bias is the tendency to search for, interpret, favor, and recall information in a way that confirms or strengthens one's prior personal beliefs or hypotheses (Nickerson 1998).

In prediction markets, confirmation bias can significantly distort the decision-making process of betting agents. Agents may seek out and give undue weight to information that aligns with their preexisting beliefs while disregarding contradictory evidence. This bias can lead to overconfidence in inaccurate predictions and the reinforcement of misguided betting strategies. As a result, the overall accuracy and efficiency of the market might be compromised, as bets are placed based on biased interpretations rather than objective analysis. Recognizing and countering confirmation bias is essential for enhancing the reliability and functionality of prediction markets as well as individual trading strategies.

Acquiescence Bias

Acquiescence bias is the tendency for respondents to agree with statements or to answer "yes" to questions, regardless of the content of the question (Paulhus 1991).

The nature of prediction markets, which often frame questions in a binary "yes" or "no" format, makes them particularly susceptible to acquiescence bias. This bias can influence the pricing of prediction market tokens, potentially leading to systematic overvaluation or undervaluation of certain outcomes. For instance, if acquiescence bias is present, participants may disproportionately choose "yes" over "no", regardless of the actual probability of the event occurring. Consequently, this creates an opportunity for informed bettors to exploit these mispricings for financial gain, particularly by betting against the prevailing bias. In essence, betting on "no" when "yes" tokens are systematically overpriced can yield profitable returns.

We can empirically test the presence of acquiescence bias in Polymarket, as the predictions are represented by “yes” and “no” tokens. Our methodology involves analyzing whether “yes” tokens are consistently overpriced relative to their true market value. The methodology and results will be presented in further sections.

Overestimating Low Probabilities

This bias refers to the tendency of individuals to overestimate the likelihood of rare events, often due to their dramatic or memorable nature (Kahneman 2011).

The tendency to overestimate low probabilities, as described by Kahneman in Thinking, Fast and Slow, is notably present in prediction markets. This bias suggests that tokens with low probabilities are underpriced, making a strategy of betting on higher-probability tokens favorable in the long run. Additionally, this indicates market inefficiency.

We can empirically test the presence of this bias in Polymarket by bundling bets and comparing the average outcomes of these bundles with the average predicted probabilities. The detailed process will be described in the following sections.

6.3 Data and Methodology

Overestimating Low Probabilities

The primary objective of this section is to investigate the correspondence between predicted probabilities and actual outcomes in prediction markets. In prediction markets, probabilities are expressed as values between 0 and 1, representing the market’s assessment of the likelihood of a particular event occurring. However, the actual outcomes are inherently binary, taking values of either 0 or 1. To bridge this discrepancy, we aggregate individual bets into bundles with varying outcomes, allowing us to calculate an average value that falls between 0 and 1. This average value can then be compared to the predicted probability, facilitating a more meaningful analysis. The bundles were generated by predicted probability deciles.

Prediction markets assign probabilities to events based on betting prices. We take a snapshot of probabilities at one point in time (for example, 90 days before the bet is evaluated). Then, we made a package of the markets by predicted probability; for example, bets with probability 0 - 10 % based on the prediction market are put into the package.

For bundles with a predicted probability between 10 % and 20 %, approximately 15 % of them should resolve as yes. If it is more than that, it could mean that it is profitable to bet on low-probability events, which would point to market inefficiency and vice versa.

The table 6.1 shows the percentage of events that occurred in each probability decile. These deciles were examined 5, 10, 30, 45, and 90 days before resolution. Often, events that are unlikely as predicted occur with even lower probabilities than the markets assigned to them.

bundle	**5**	**10**	**30**	**45**	**90**
0.0-0.1	0.0%	0.9%	1.1%	1.3%	2.4%
0.1-0.2	9.5%	4.0%	0.0%	6.3%	10.5%
0.2-0.3	12.5%	7.1%	16.7%	21.1%	20.0%
0.3-0.4	22.2%	16.7%	9.1%	33.3%	18.8%
0.4-0.5	48.0%	47.4%	48.2%	49.1%	50.6%
0.5-0.6	52.9%	58.3%	58.8%	55.0%	43.5%
0.6-0.7	81.8%	83.3%	91.7%	70.0%	86.7%
0.7-0.8	85.7%	92.9%	83.3%	80.0%	82.4%
0.8-0.9	91.3%	96.6%	100.0%	88.2%	89.5%
0.9-1.0	100.0%	99.1%	98.8%	100.0%	97.5%

Table 6.1: Predicted Probabilities for Decile Bundles: Average Predicted Probability for Each Bundle 5, 10, 30, 45, and 90 Days in Advance

Based on Kahneman's research Kahneman (2011), people tend to overestimate low probabilities. We can try to confirm or refute this in the following exploration.

Modeling

For modeling purposes, the markets were divided into bundles based on predicted probabilities 90 days before resolution, with intervals of 0.05, resulting in 20 distinct points for the linear regression analysis. Each bundle represents a range of predicted probabilities. For each bundle, we calculated the average output, the average predicted probability of the outcome, and the count of markets in the given bundle.

To ensure objectivity, it is crucial to weigh each bundle according to the number of observations it contains. This approach ensures that bundles with more observations have a proportionately larger impact on the regression analysis, thus providing a more accurate representation of the data.

We employed a linear regression model with the average output as the dependent variable and the average prior probability of outcome as the independent variable. Given that the number of observations in each bundle varies, we utilized a Weighted Least Squares (WLS) regression model. In this model, the weights are determined by the count of markets in each bundle, thereby incorporating the varying number of data representations across different bundles.

The expected value of the β_1 coefficient in this model is 1, reflecting the assumption that the predicted probabilities should match the observed final probabilities. This alignment would indicate that the prediction market is well-calibrated.

To determine if the β_1 coefficient significantly deviates from the expected value of 1, we conducted a hypothesis test. The specific hypotheses tested are following: The null hypothesis H_0 and the alternative hypothesis H_1 can be written as:

$$H_0 : \beta_1 = 1$$

$$H_1 : \beta_1 > 1$$

The hypothesis test employed is a t-test, evaluating whether the β_1 coefficient is statistically significantly greater than 1. The rationale behind this hypothesis is that a β_1 coefficient greater than 1 would imply that the market tends to overestimate low and underestimate high probabilities.

Based on the Ramsey test, we verified that the model is likely misspecified. The author has thus prepared a monotone function to represent the overestimation bias of low-probability events. The function is designed to satisfy the following properties:

1. **Defined on the interval [0,1]**: This is essential as we are dealing with probabilities, and probability is, by definition, on this interval.

2. **Monotonicity**: The function is monotonic. Even in the presence of low-probability overestimation, it is not expected that any specific lower probability level would be overestimated to a higher level than more probable ones.

3. **Value at 0.5**: The function is defined such that $f(0.5) = 0.5$. This is because 0.5 represents neither a high nor a low probability, indicating that there should be no bias at this midpoint.

4. **Central Symmetry around [0.5:0.5]**: The function is centrally symmetric around the point 0.5. This symmetry implies that overestimating small probabilities is equivalent to underestimating large probabilities by the same magnitude. This is due to the fact that for every event with a small probability of occurrence, there is an opposite event with a high probability of non-occurrence, and the sum of these probabilities is always 1.

5. **Continuity**: The function is continuous, wee no reason to expect the bias to change abruptly.

6. **Convexity on [0, 0.5]**: The function is convex in the interval [0, 0.5], indicating that as probabilities increase towards 0.5, the bias decreases at an increasing rate. By the property of central symmetry, the function is concave on the interval [0.5, 1], indicating a similar behavior for high probabilities.

The proposed prescription function is as follows:

$$\sqrt{|x - 0.5|} \cdot \mathrm{sgn}(x - 0.5) \cdot \sqrt{\frac{1}{2}} + \frac{1}{2}$$

which is plotted in figure 6.1.

Thus, we created a model that uses this function to estimate prediction markets. Using this model, the Ramsey test did not show that the model was misspecified. The main criterion for evaluating this model is adjusted R squared since it is not possible to interpret this model based on the coefficient due to the properties of the function used.

Acquiescence Bias

Acquiescence bias, the tendency of respondents to agree with statements regardless of their content, can significantly distort market predictions, particularly in the context of prediction markets where participants place bets on the likelihood of future events. This study investigates the presence and impact of acquiescence bias in the Polymarket prediction market. By analyzing market

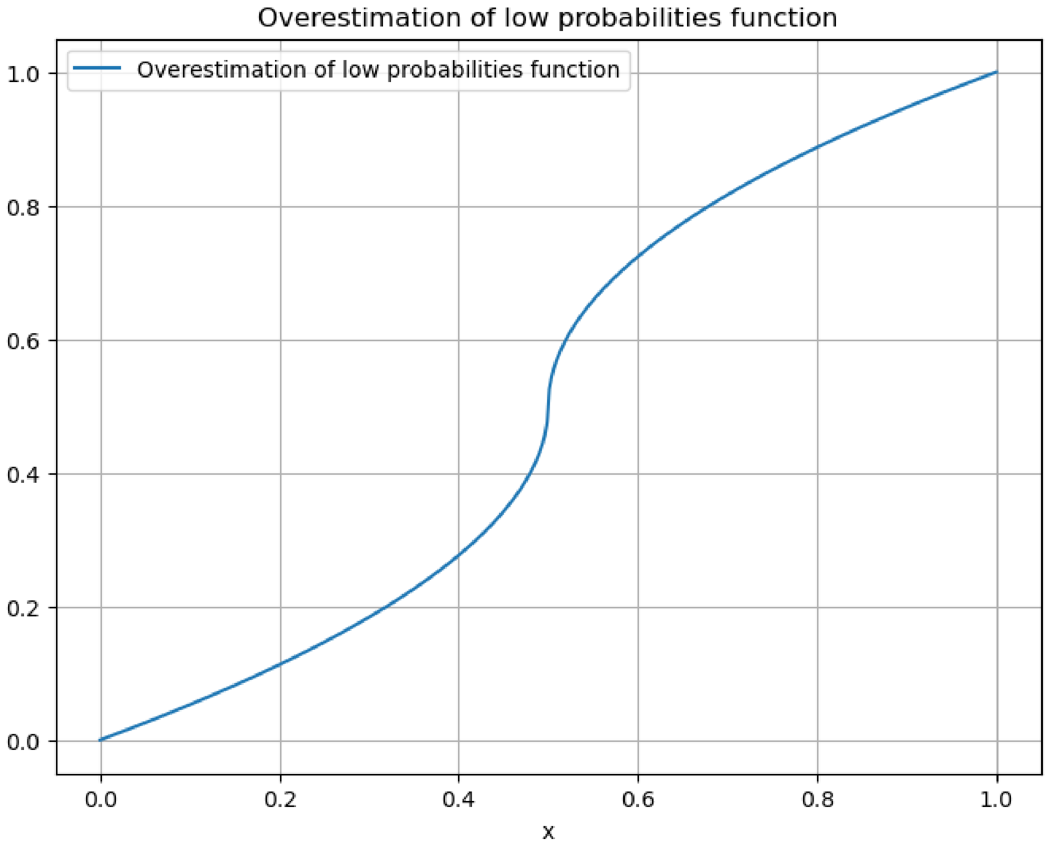

Figure 6.1: Overestimation of low probabilities function

data snapshots taken at various intervals before the resolution of bets, we aim to quantify this bias and understand its temporal dynamics. Our methodology involves converting token prices into implied probabilities, comparing expected and actual outcomes, and employing statistical tests to validate our findings.

We destined market data snapshots at four critical time intervals: 90, 60, 30, and 15 days before the resolution of each market bet. For each snapshot and each market, we calculated the implied probability of a "YES" outcome based on the token prices. This involved converting market prices into probabilities, which reflect the collective expectation of the event occurring. We aggregated data across all examined markets, and we calculated the expected number of "yes" outcomes as the sum of the individual market's expected "YES" outcomes. This provided a benchmark to compare against actual market outcomes.

To rigorously test our hypothesis, we employed a binomial test. The null hypothesis (H0) posited that the proportion of actual "yes" outcomes equals the market's expected proportion. The alternative hypothesis (H1) suggested that the market's expected proportion of "yes" outcomes exceeds the actual

Parameter	Coefficient	SE	T-stat	P-value
const	-0.06	0.04	-1.80	0.09
average_prior_p_outcome	1.13	0.0580	19.56	0.00***
R-squared:		0.955		
Adj. R-squared:		0.953		

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 6.2: Overestimating Low Probabilities: WLS Regression Results

proportion.

$$H_0 : \text{expected yes outcome} = \text{yes outcome}$$

$$H_1 : \text{expected yes outcome} > \text{yes outcome}$$

This methodological framework allows us to determine whether acquiescence bias significantly impacts the pricing of "yes" tokens in prediction markets. If "yes" tokens are found to be consistently overpriced, it would indicate the presence of this bias, validating the hypothesis that prediction market participants disproportionately favor affirmative outcomes. This finding would not only contribute to the academic understanding of behavioral biases in financial markets but also offer practical insights for market participants seeking to optimize their betting strategies.

6.4 Results and Discussion

Overestimating Low Probabilities

The Weighted Least Squares (WLS) regression results are presented in Table 6.2. The model exhibits an excellent fit, with an R-squared value of 0.955, indicating that approximately 95.5 % of the variability in the average output is accounted for by the average prior probability of the outcome. To assess the presence of Overestimating Low Probabilities Bias, we examine whether β_1 equals 1. If β_1 is significantly greater than 1, this suggests a statistically significant bias.

Parameter	Coefficient	SE	P-value
const	-0.0181	0.021	0.410
average_prior_p_outcome_sqrt2	1.0438	0.034	0.000***
R-squared:		0.981	
Adj. R-squared:		0.980	

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 6.3: Overestimating Low Probabilities: WLS Regression Results Using Low Probability Overestimation Bias Function

$$\widehat{\beta}_1 = 1.13$$
$$\beta_1 = 1$$
$$\mathrm{SE}(\widehat{\beta}_1) = 0.058$$

The test statistic t is calculated as:

$$t = \frac{\widehat{\beta_1} - \beta_1}{\mathrm{SE}(\widehat{\beta_1})} = \frac{1.13 - 1}{0.058} = \frac{0.13}{0.058} \approx 2.18$$

To determine the p-value, we compare the test statistic $t = 2.1836$ with the critical value from the t-distribution with 18 degrees of freedom for a one-tailed test at $\alpha = 0.05$:

$$\text{Critical value } (t_{0.05,18} = 1.734)$$

Since $2.1836 > 1.734$, we reject the null hypothesis H_0.

Therefore, there is sufficient evidence to support the claim that the coefficient for “average_prior_p_outcome” is greater than 1. Based on this model, we can confirm the presence of Overestimation of low probabilities on the poly market.

As detailed in the methodology section, the initial model likely did not fulfill all necessary requirements. Consequently, we developed an alternative approach to test for low probability overestimation within its framework. This new model incorporated a specific function designed to address the bias of low probability overestimation. The results of this model are presented in Table 6.3.

The findings indicate that the relationship between predicted and actual probabilities is not linear. The function included in the model accurately captures this relationship. According to the Ramsey test, the model is well-specified and meets all necessary assumptions, achieving a high R-squared value of 98 %. The model line with the data is plotted in figure 6.2.

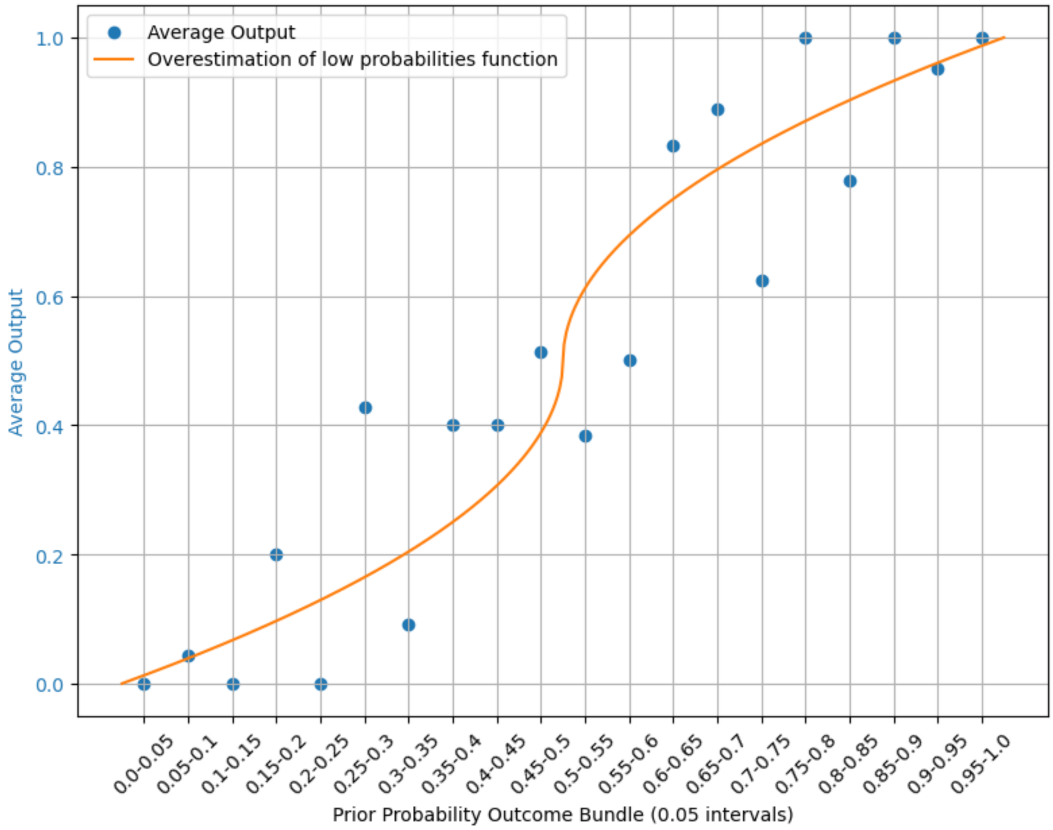

Figure 6.2: Overestimation of low probabilities function with data

The analysis shows that the polymarket exhibits a bias, overestimating low probabilities and consequently underestimating high probabilities. Based on this, traders should focus on high-probability bets. However, it is important to note that this bias may be temporary. As demonstrated in the case of corporate prediction markets, such biases tend to diminish over time as the market matures Cowgill & Zitzewitz (2013).

Acquiescence Bias

The analysis of acquiescence bias in Polymarket prediction markets yielded the following table 6.3:

The results show that statistically significant acquiescence bias ($\alpha = 0.05$)

Days in advance	Expected Yes	Actual Yes	p-value
90	72.5	44	$\ll 0.001$
60	67.8	44	$\ll 0.001$
30	58.3	44	0.03
15	49.5	44	0.41

Table 6.4: Expected vs. Actual 'Yes' Outcomes: Comparison at Different Days in Advance

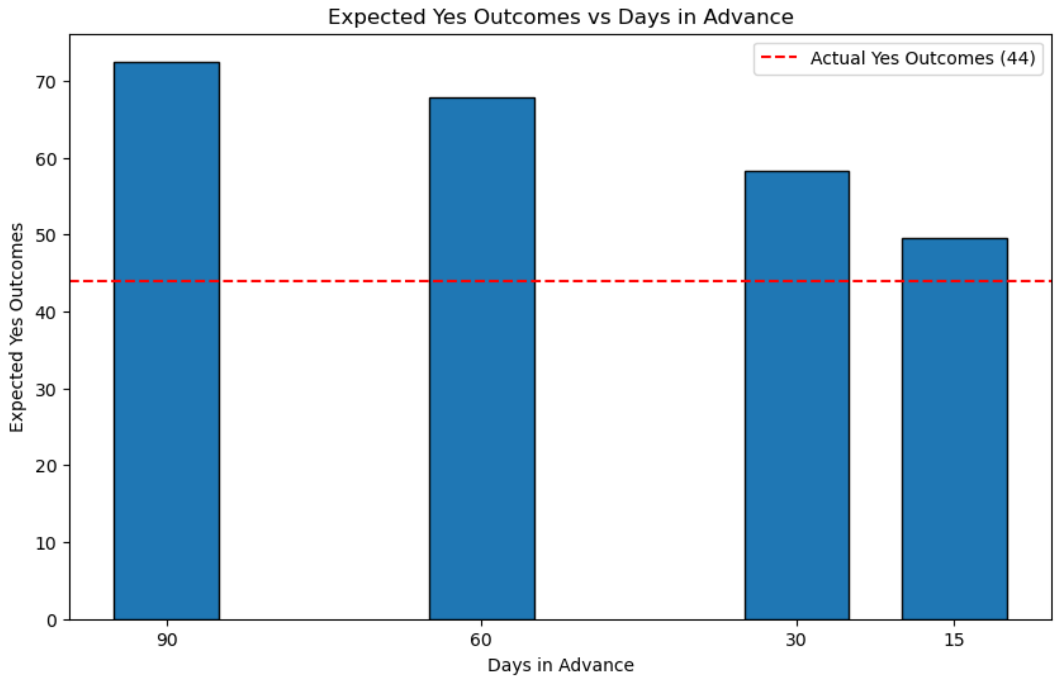

Figure 6.3: Expected 'Yes' Outcomes 90, 60, 30, and 15 Days in Advance: Comparison to Actual 'Yes' Outcomes

exists at 90, 60, and 30 days before the resolution of the bets. The p-values for these time intervals are significantly below the 0.05 threshold, indicating strong evidence against the null hypothesis. However, 15 days before the bet resolution, the p-value of 0.41 suggests that the bias is no longer statistically significant, implying a convergence of market expectations with time to the actual outcomes.

The observed results reveal several key insights. Temporal Influence of Bias: The presence of acquiescence bias is more pronounced at 90, 60, and 30 days before the resolution of the bets. As the time to the resolution decreases, the bias diminishes, indicating that the markets become more accurate in their predictions as they converge toward the final outcome. This temporal pattern suggests that initial market prices may be influenced by biases, which are gradually corrected as more information becomes available. Market Convergence: The diminishing effect of acquiescence bias as the resolution date nears is a testament to the self-correcting nature of prediction markets. As more information is assimilated, market participants adjust their expectations, leading to a convergence of the market price to the true and final probability of the event. Bias Magnitude and Overestimation: Despite the general presence of acquiescence bias, it is important to note that the overall likelihood of "YES" outcomes in Polymarket bets is relatively low. Only 44 out of the 187 markets examined resulted in a "YES" outcome. This low rate suggests that part of the observed bias might be due to a general tendency to overestimate low-probability events rather than acquiescence bias alone. This distinction is crucial for understanding the underlying dynamics driving the market prices and developing more refined betting strategies.

The empirical investigation into Polymarket prediction markets has highlighted the presence of acquiescence bias, particularly at longer time horizons before the resolution of bets. This bias tends to diminish as the resolution date approaches, demonstrating the market's ability to self-correct and align with actual probabilities. The findings underscore the importance of considering both behavioral biases and the inherent probability structure of events when analyzing and participating in prediction markets. Future research could further explore the interaction between different biases and market mechanisms, enhancing our understanding of prediction market dynamics.

6.5 Conclusion

We investigated the presence and impact of acquiescence bias and the overestimation of low probabilities in the Polymarket prediction market. Our analysis reveals significant insights into the behavioral biases influencing market predictions and provides evidence of the market's ability to self-correct over time.

Firstly, the Weighted Least Squares (WLS) linear regression results showed that the coefficient for "average_prior_p_outcome" is significantly greater than 1, confirming the overestimation of low probabilities.

The alternative model developed to address low probability overestimation further supported this finding. The model incorporated a special function to better model the bias and achieved a high R-squared value of 98 %. This suggests that the relationship between predicted and actual probabilities is not linear and that the overestimation of low probabilities is a significant factor in market pricing dynamics.

Additionally, the examination of acquiescence bias through temporal analysis of market data snapshots revealed that this bias is more pronounced at longer time intervals before the resolution of bets (90, 60, and 30 days). The bias diminishes as the resolution date approaches, indicating a convergence of market expectations with actual outcomes. This temporal influence suggests that initial market prices may be influenced by biases, which are gradually corrected as more information becomes available.

The empirical findings demonstrate that Polymarket tends to overestimate low-probability events and that acquiescence bias significantly impacts market predictions, particularly at longer time horizons. However, as the market approaches the resolution date, these biases diminish, showcasing the market's ability to self-correct and align with actual probabilities.

These insights have important implications for market participants and researchers. For traders, understanding the presence of these biases can inform more strategic betting decisions, particularly by focusing on high-probability bets as the resolution date nears. For researchers, the study underscores the importance of considering both behavioral biases and the inherent probability structure of events when analyzing prediction markets.

Chapter 7

Conclusion

This thesis explores the dynamics of prediction markets within the cryptocurrency landscape, particularly focusing on the Polymarket. We have delivered a deeper understanding of this emerging financial instrument by providing a detailed analysis of their functionality, examining the potential for predicting future events, and measuring volatility.

Our investigation into the efficiency of prediction markets underscores the importance of your work in this field. While these markets offer valuable insights and aggregate information in time, they are still influenced by various cognitive biases. Overestimating low probabilities and acquiescence bias are notably present, impacting market predictions. This highlights the need for your continued refinement and maturity in these markets.

The volatility analysis in Polymarket compared to traditional financial instruments like stocks and cryptocurrencies reveals significantly higher risk associated with prediction markets. This heightened volatility requires traders to adopt more sophisticated risk management strategies to navigate these markets effectively.

characterized by the overestimation of affirmative outcomes and the overestimation of low probabilities. Empirical evidence supports the existence of both biases in Polymarket. The presence of biases such as overconfidence, loss aversion, gambler's fallacy, sunk cost fallacy, and confirmation bias should be examined and measured in further research.

Overall, this thesis contributes to the broader field of behavioral finance and enhances our understanding of prediction markets in the cryptocurrency ecosystem. By recognizing and addressing the biases and volatility in these markets, traders can make more informed decisions, ultimately leading to more efficient and accurate market predictions. This research has direct implications for the work of traders and analysts in the field, making it highly relevant and applicable.

Future research should explore the interplay between various biases and Polymarket more thoroughly. Empirical studies are needed to investigate the presence and impact of biases such as overconfidence, loss aversion, the Gambler's fallacy, the sunk cost fallacy, and confirmation bias. Further research might help the traders, leading them to increase the accuracy of Polymarket probability prediction.

Bibliography

Adams, H., N. Zinsmeister, D. Robinson, & M. Salem (2019): "Uniswap: A simple decentralized exchange." *arXiv preprint arXiv:1901.01253* .

Adams, H., N. Zinsmeister, D. Robinson, & M. Salem (2020): "Uniswap v2 core." `https://uniswap.org/whitepaper.pdf`.

Angeris, G. & T. Chitra (2020): "An analysis of uniswap markets." *arXiv preprint arXiv:2005.04645* .

Archak, N. & P. G. Ipeirotis (2009): "Modeling Volatility in Prediction Markets." *EC'09 Proceedings* .

Arkes, H. R. & C. Blumer (1985): "The Psychology of Sunk Cost." *Organizational Behavior and Human Decision Processes* **35(1)**: pp. 124–140.

Arrow, K. J., R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. O. Ledyard, S. Levmore, R. E. Litan, P. Milgrom, F. D. Nelson *et al.* (2008): "The promise of prediction markets." *Science* **320(5878)**: pp. 877–878.

Atkinson, R. (2012): "Acquiescence response bias in market research." *International Journal of Market Research* **54(3)**: pp. 315–337.

Barberis, N. C. (2013): "Thirty years of prospect theory in economics: A review and assessment." *Journal of Economic Perspectives* **27(1)**: pp. 173–196.

Berentsen, A. & F. Schar (2019): *Stablecoins: The Quest for a Low-Volatility Cryptocurrency*, pp. 65–71. London: CEPR Press.

Berg, J. E., R. Forsythe, F. D. Nelson, & T. A. Rietz (2008): "Prediction markets for economic forecasting." *International Journal of Forecasting* **24(2)**: pp. 283–298.

Berg, J. E., F. D. Nelson, & T. A. Rietz (2003): "Accuracy and forecast standard error of prediction markets." *Department of Economics, University of Iowa* **26**.

Berg, J. E. & T. A. Rietz (2003): "Prediction Markets as Decision Support Systems." *Information Systems Frontiers* **5(1)**: pp. 79–93.

Bjorvatn, K. & C. Eckel (2006): "Policy Competition for Foreign Direct Investment Between Asymmetric Countries." *European Economic Review* **50(7)**: pp. 1891–1907.

Blomstrom, M. & A. Kokko (2003): "The Economics of Foreign Direct Investment Incentives." *NBER Working Papers 9489*, National Bureau of Economic Research, Inc.

Buckley, P. (2022): "Blockchain Based Prediction Markets." *UK Academy for Information Systems Conference Proceedings 2022* .

Buterin, V. (2013): "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform."

Caldarelli, G. (2020): "Understanding the Blockchain Oracle Problem: A Call for Action." *Information* .

Celik, L. & E. Onur (2010): "Determination of Odds in Prediction Markets: Coexistence of Posted-offer and Double-auction Designs." *CERGE-EI Repository* .

Consortium, C. (2021): "Usdc: Us dollar coin." *Centre* `https://www.centre.io/usdc`.

Cont, R. (2001): "Empirical properties of asset returns: stylized facts and statistical issues." *Quantitative Finance* **1(2)**: pp. 223–236.

Cowgill, B., J. Wolfers, & E. Zitzewitz (2009a): "Using prediction markets to track information flows: Evidence from google." Manuscript, The Wharton School, University of Pennsylvania.

Cowgill, B., J. Wolfers, & E. Zitzewitz (2009b): "Using Prediction Markets to Track Information Flows: Evidence from Google." *Unknown* .

Cowgill, B. & E. Zitzewitz (2013): "Corporate Prediction Markets: Evidence from Google, Ford, and Koch Industries." *Unknown* .

Cowgill, B. & E. Zitzewitz (2015): "Corporate Prediction Markets: Evidence from Google, Ford, and Firm X." *Review of Economic Studies* **82**: pp. 1309–1341.

Dardaman, E. & A. Gupta (2023): "Asking Better Questions: The Art and Science of Forecasting." *ACM* .

Djankov, S. & P. Murrell (2002): "Enterprise Restructuring in Transition: A Quantitative Survey." *Journal of Economic Literature* **40(3)**: pp. 739–792.

Eisenberg, R. & A. Lehar (2020): "Smoothing mechanisms for price stability in decentralized finance." *arXiv preprint arXiv:2012.08873* .

Fama, E. F. (1970): "Efficient capital markets: A review of theory and empirical work." *Journal of Finance* .

Feng, Z., M. Hougan, & D. Lawant (2021): "Composability of decentralized finance protocols." *arXiv preprint arXiv:2102.07718* .

Forsythe, R., F. Nelson, G. R. Neumann, & J. Wright (1992): "Anatomy of an experimental political stock market." *The American Economic Review* **82(5)**: pp. 1142–1161.

Gjerstad, S. (2005): "Risk Aversion, Beliefs, and Prediction Market Equilibrium." *Economic Science Laboratory, McClelland Hall 116, University of Arizona* .

Gonzalez, R. & G. Wu (1999): "On the shape of the probability weighting function." *Cognitive Psychology* **38(1)**: pp. 129–166.

Haaparanta, P. (1996): "Competition for Foreign Direct Investment." *Journal of Public Economics* **63(1)**: pp. 141–53.

Hanson, R. (2003): "Combinatorial information market design." *Information Systems Frontiers* **5(1)**: pp. 107–119.

Hanson, R. (2006): "Information aggregation and manipulation in an experimental market." *Journal of Economic Behavior & Organization* **60(4)**: pp. 449–459.

Hanson, R. (2007): "Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation." *The Journal of Prediction Markets* **1**: pp. 3–15.

Harz, D. & S. Bano (2021): "A short overview of decentralized prediction markets." *arXiv preprint arXiv:1907.03977* .

Haufler, A. & I. Wooton (2006): "The Effects of Regional Tax and Subsidy Coordination on Foreign Direct Investment." *European Economic Review* **50(2)**: pp. 285–305.

Heimbach, L., Y. Wang, & R. Wattenhofer (2023): "Behavior of Liquidity Providers in Decentralized Exchanges." *ETH Zurich* .

Kahneman, D. (2011): *Thinking, Fast and Slow.* Farrar, Straus and Giroux.

Kahneman, D. & A. Tversky (1979a): "Prospect theory: An analysis of decision under risk." *Econometrica* **47(2)**: pp. 263–291.

Kahneman, D. & A. Tversky (1979b): "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* **47(2)**: pp. 263–291.

Kapp-Schwoerer, L. (2023): "Improved Liquidity for Prediction Markets." *Distributed Computing Group Computer Engineering and Networks Laboratory, ETH Zürich* .

Knowles, E. S. & J. A. Nathan (2005): "Acquiescence bias in survey research." *Encyclopedia of Social Measurement* **1**: pp. 1–7.

Lewis, K. K. (2011): "Market expectations and macroeconomic outcomes." *Journal of Monetary Economics* **58(6-8)**: pp. 594–611.

Lipton, A. & T. Hardjono (2020): "Stablecoins: risks, benefits and opportunities." *The Journal of Risk Finance* .

Manski, C. F. (2004): "Measuring Expectations." *Econometrica* **72(5)**: pp. 1329–1376.

Manski, C. F. (2006a): "Interpreting the predictions of prediction markets." *Economics Letters* **91(3)**: pp. 425–429.

Manski, C. F. (2006b): "Interpreting the Predictions of Prediction Markets." *Unknown* .

Nakamoto, S. (2008): "Bitcoin: A Peer-to-Peer Electronic Cash System." Https://bitcoin.org/bitcoin.pdf.

Nickerson, R. S. (1998): "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology* **2(2)**: pp. 175–220.

Oettler, M. (2021): "Types of Prediction Markets."

Ottaviani, M. & P. N. Sorensen (2009): "The wisdom of the few." *Journal of Economic Theory* **144(5)**: pp. 2035–2058.

Paulhus, D. L. (1991): "Measurement and Control of Response Bias." In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (editors), "Measures of Personality and Social Psychological Attitudes," pp. 17–59. Academic Press.

Polymarket (2023a): "Polymarket documentation." `https://polymarket.com/documentation`.

Polymarket (2023b): "Polymarket github repository." `https://github.com/Polymarket`.

Pompian, M. M. (2006): *Behavioral Finance and Wealth Management: How to Build Optimal Portfolios That Account for Investor Biases.* John Wiley Sons.

Ritholtz, B. (2008): *The importance of liquidity in financial markets.* Wiley.

Rothschild, D. (2009): "Forecasting elections: Comparing prediction markets, polls, and their biases." *Public Opinion Quarterly* **73(5)**: pp. 895–916.

Slamka, C., B. Skiera, & M. Spann (2013): "Prediction Market Performance and Market Liquidity: A Comparison of Automated Market Makers." *IEEE Transactions on Engineering Management* **60(1)**: p. 169.

Snowberg, E., J. Wolfers, & E. Zitzewitz (2012): "Prediction markets for economic forecasting." *American Economic Review* **102(4)**: pp. 1–36.

Spann, M. & B. Skiera (2003): "Internet-based virtual stock markets for business forecasting." *Management Science* **49(10)**: pp. 1310–1326.

Spann, M. & B. Skiera (2009): "Sports forecasting: A comparison of the forecast accuracy of prediction markets, betting odds and tipsters." *Journal of Forecasting* **28(1)**: pp. 55–72.

Stanley, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

SUROWIECKI, J. (2005): *The Wisdom of Crowds.* Anchor.

TAPSCOTT, D. & A. TAPSCOTT (2017): "How blockchain is changing finance." *Harvard Business Review* **1(9)**: pp. 2–5.

TEAM, P. R. (2021): "Volatility patterns in prediction markets: A comparative analysis." *Polymarket Research Journal* .

TETLOCK, P. C., M. SAAR-TSECHANSKY, & S. MACSKASSY (2008): "More than words: Quantifying language to measure firms' fundamentals." *The Journal of Finance* **63(3)**: pp. 1437–1467.

TETLOCK, P. E. & B. MELLERS (2015): *Superforecasting: The art and science of prediction.* Crown.

TVERSKY, A. & D. KAHNEMAN (1971): "Belief in the Law of Small Numbers." *Psychological Bulletin* **76(2)**: pp. 105–110.

TZIRALIS, G. & I. TATSIOPOULOS (2007): "Prediction Markets: An Extended Literature Review." *The Journal of Prediction Markets* **1**: pp. 75–91.

UNDERWOOD, S. (2016): "Blockchain beyond bitcoin." *Communications of the ACM* **59(11)**: pp. 15–17.

URQUHART, A. (2016): "The inefficiency of bitcoin." *Economics Letters* **148**: pp. 80–82.

WELLS, L. T., N. ALLEN, J. MORISSET, & N. PIRNIA (2001): *Using Tax Incentives to Compete for Foreign Investment: Are They Worth the Cost?* Washington, DC: FIAS.

WERNER, S. M., R. AUER, & G. VON PETER (2021): "Decentralized finance: On blockchain- and smart contract-based financial markets." *FSI Insights on Policy Implementation* **(37)**.

WOLFERS, J. & E. ZITZEWITZ (2004a): "Prediction Markets." *Journal of Economic Perspectives* **18(2)**.

WOLFERS, J. & E. ZITZEWITZ (2004b): "Prediction markets." *Journal of Economic Perspectives* **18(2)**: pp. 107–126.

WOLFERS, J. & E. ZITZEWITZ (2006): "Interpreting Prediction Market Prices as Probabilities." *NBER Working Paper Series* **(12200)**.

Xu, X., I. Weber, & M. Staples (2021): "A systematic review of blockchain." *Communications of the ACM* **63(10)**: pp. 68–78.

Zhan, J., C. W. Lau, L. Rahim, & Q. Le (2021): "Decentralized prediction markets." .

Zhang, S. K., I. Bentov, M. Nagler, & G. Neu (2020): "A solution to the oracle problem using peer prediction and decentralized information markets." *Cryptoeconomic Systems* .